Paper 206-2012

Finding the Winning Combination: An Application of Multivariate Testing from Digital Marketing

Robert Krutsick

ABSTRACT

When designing a marketing campaign, marketers must identify the appropriate mix of offer elements to optimize their desired measure of campaign performance (such as response, conversion, sales, and profit). This is not usually a trivial task since the number of offer elements making up a typical marketing campaign can be quite large. The chosen mix of offer elements included in a campaign is usually based on a combination of business judgment and the results of experimental testing. In this paper, we present an application of multivariate testing from the field of digital marketing analytics. We demonstrate how to design a fractional factorial test with JMP® and how to predict the optimal combination of offer elements with the GLIMMIX procedure.

INTRODUCTION

A common challenge faced by marketers is determining the right combination of offer elements to include in a targeted marketing campaign. Typical offer elements include ad copy, creative images, price, and promotional incentives. Ideally, the values of these offer elements are selected to maximize some measure of expected campaign performance. One way to determine those optimal values is by conducting marketing experiments using A/B or multivariate testing. In some testing scenarios, the number of factors to consider is quite large, making it infeasible or prohibitively expensive to test all possible combinations in a full factorial design. For example, a full factorial design with 10 factors each at two levels requires $2^{10} = 1,024$ different test cells. Clearly it is not feasible to run a marketing test that includes 1,024 test cells.

Fortunately, fractional factorials enable marketers to test one-half, one-fourth, or even smaller fractions of the number of cells required by a full factorial design. In a full factorial design, all main effects and interactions among the factors can be estimated. The fact that all of these effects can be estimated does not mean that they are all important. In many situations, it can safely be assumed that the main effects are larger than the two-way interactions, which are in turn larger than the three-way interactions, and so on. Fractional factorial designs exploit this fact and allow marketers to execute smaller tests at the expense of giving up the ability to estimate some of the higher order interactions. In many scenarios this is a reasonable tradeoff. In this paper we demonstrate the use of SAS® software and JMP to design and analyze a fractional factorial test from an email marketing campaign.

THE EXPERIMENT

An online retailer wanted to conduct an email marketing campaign to reactivate inactive customers who had not purchased from the web store within the past 12 months. The goal was to maximize the conversion rate of the campaign. Conversion rate is defined as the total number of purchases divided by the total number of email sends. Before launching the full campaign, the retailer wanted to conduct an experiment to identify the combination of offer elements that would yield the highest predicted conversion rate.

Table 1 describes the six factors (offer elements) included in the test. Each factor was tested at two levels. Level 1 represented the current form of the offer and level 2 represented the new level to be tested. In the data file, level 1 was coded as 0 and level 2 was coded as 1. The outcome variable was coded as 1 if the offer resulted in a purchase and was coded as 0 otherwise.

| Factor	Level 1	Level 2
A) Personalization	Dear Valued Customer,	Dear Your Name,
B) Graphic	Standard	Enhanced
C) Message	We've Missed You	Come See What's New
D) Copy	Standard	New
E) Shipping and Handling	Standard Shipping Rates	Free Shipping
F) Incentive	$10 off order of $100	$15 off order of $150

Table 1. Factors and Factor Levels Included in the Test.

A full factorial design of this test would require $2^6 = 64$ test cells, so a more efficient design was needed in order to reduce cost without losing the ability to estimate the main effects and interactions assumed to be most important. A

2^{6-2} fractional factorial design was selected (also known as a one-quarter or one-fourth fractional factorial). This design required only $2^{6-2} = 16$ test cells and provided sufficient resolution to estimate the effects deemed to be most important based on domain knowledge.

The fractional factorial test design was created with JMP 8.02 using the following steps.

1.   Select **DOE >Screening Design**.
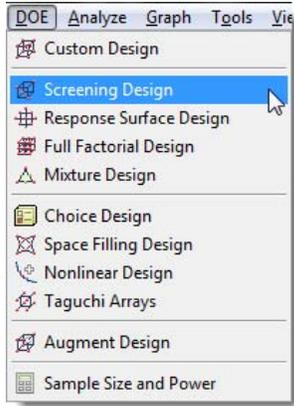
Display 1. DOE Menu.

2.   In the Responses panel, enter the response variable name and goal. In the Factors panel, enter six 2-level categorical factors and click add. Type the name of each factor and specify its role. After all information has been entered, click **Continue**.
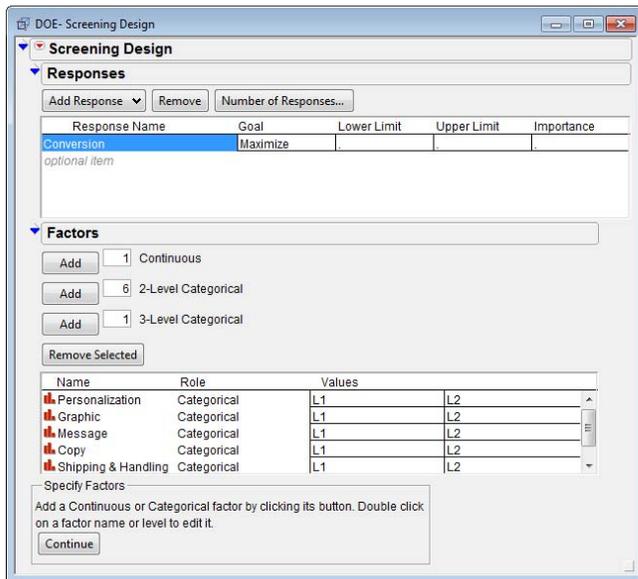
Display 2. Responses and Factors.

3.   From the Design List panel, highlight the fractional factorial design corresponding to 16 runs with no blocking (Block Size is blank) then click **Continue**.
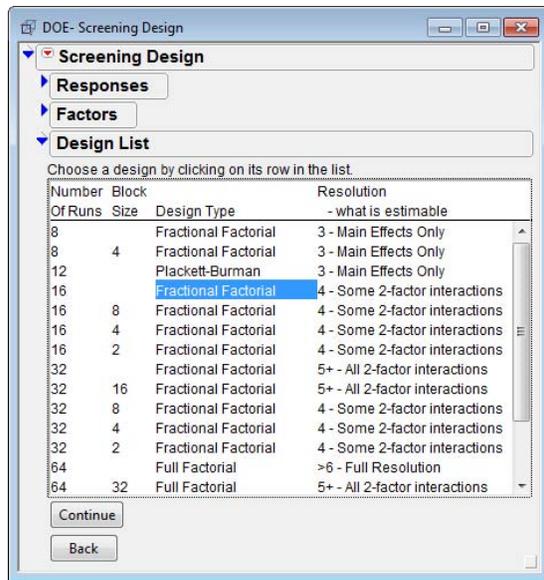
Display 3. Design List with Fractional Factorial Design Selected.

4. Examine the Aliasing of Effects panel. This panel shows which effects are confounded with other effects in the chosen design. When designing a fractional factorial test, it is important to review the aliasing of effects to be sure all effects that are expected to important can be estimated unambiguously. Care was taken to choose a design that did not mask any main effects or second order interactions that were expected to be important. For more details about design resolution and aliasing in fractional factorials, see Box, Hunter, and Hunter (1978), Kuehl (2000), or Ledolter and Swersey (2007).
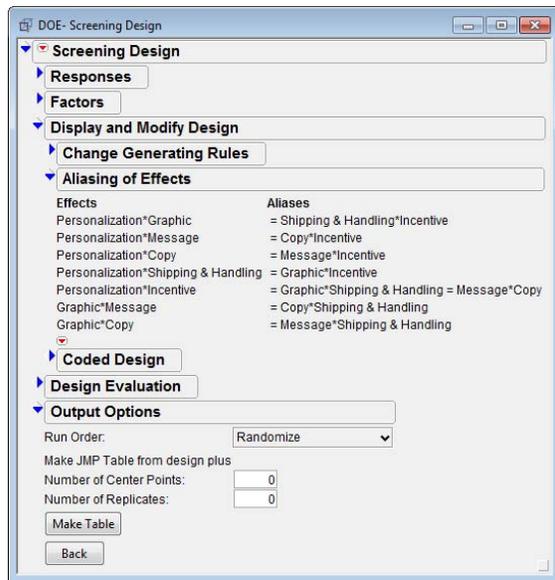
Display 4. Aliasing of Effects Panel.

5. When satisfied with the aliasing of effects, click **Make Table** to produce the design table for the experiment. Display 5 shows the design table generated for the chosen experiment. This table shows the 16 test cells in the fractional factorial design. Each row in the table contains instructions for producing one of the email versions required for the test. For example, the first row specifies that Personalization, Graphic, Shipping and Handling, and Incentive should be set to L1 (level 1), while Message and Copy should be set to L2 (level 2). Using this table as a guide, an email version was created for each of the 16 test cell recipes.

3

Display 5. Design Table for the Experiment.

After the design table and email versions were created, 32,000 customers were randomly selected from the inactive file and divided into 16 groups of 2,000. Each group of 2,000 customers was randomly assigned to one of the 16 test cells. The email offers were sent to the customers and within two weeks final results were available for analysis.

TEST RESULTS

The final results of the test campaign were compiled and stored in a data set named 'email'. The first 15 observations from the data set are shown in Output 1. Each row corresponds to a customer who received one of the email offers. The variable 'ID' is a unique customer identifier. The variable 'Test Cell' identifies the email version that the customer received where the values correspond to one of the 16 rows in the design table. The next six variables identify the levels of each experimental variable included in the design. The last variable 'Conversion' identifies whether the customer responded to the campaign and made a purchase. Of the 32,000 customers who received the email offer, 990 converted (3.09%).

```
First 15 Observations in Data Set: Email

     Test
ID   Cell   Personal   Graphic   Message   Copy   Shipping   Incentive   Conversion


1    1         0          0         1        1        0          0            0
2    1         0          0         1        1        0          0            0
3    1         0          0         1        1        0          0            0
4    1         0          0         1        1        0          0            0
5    1         0          0         1        1        0          0            0
6    1         0          0         1        1        0          0            0
7    1         0          0         1        1        0          0            0
8    1         0          0         1        1        0          0            0
9    1         0          0         1        1        0          0            0
10   1         0          0         1        1        0          0            0
11   1         0          0         1        1        0          0            0
12   1         0          0         1        1        0          0            0
13   1         0          0         1        1        0          0            0
14   1         0          0         1        1        0          0            0
15   1         0          0         1        1        0          0            0
```

Output 1. Printout of the First 15 Observations in the Email Data Set.

THE MODEL

The outcome variable in the experiment is binary so logistic regression is an appropriate method for analyzing the relationship between the outcome and the experimental factors. Let p_i be the probability of conversion and let $Personal_i, Graphic_i, Message_i, Copy_i, Shipping_i,$ and $Incentive_i$ be the experimental indicator variables for the ith subject. The model contains all six main effects and a subset of the two-way interactions thought to be most

important based on domain knowledge. The remaining interactions are not included in the model because they were expected to be negligible.

$$\log\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1(Personal_i) + \beta_2(Graphic_i) + \beta_3(Message_i) + \beta_4(Copy_i) + \beta_5(Shipping_i) + \beta_6(Incentive_i)$$
$$+ \beta_7(Shipping_i * Incentive_i) + \beta_8(Personal_i * Message_i)$$

Even though our experiment included only a fraction of the cells in a full factorial design, we were able to use our results to estimate all the parameters in the above model. We then used the fitted model to obtain conversion rate predictions for all 64 combinations of our experimental variables, even the combinations that were not included in our test. We used the following PROC GLIMMIX step to obtain the parameter estimates.

```
proc glimmix data=email;
model conversion(event='1') = personal graphic message copy shipping incentive
shipping*incentive personal*message / dist=binary link=logit s;
run;
```

The parameter estimates for the fitted model are shown in Output 2. The variables Message, Shipping, Incentive, Graphic, and Shipping*Incentive are all significant at the 10% level. The parameter estimate on the Shipping*Incentive interaction is negative and lower in magnitude than the parameter estimates on the Shipping and Incentive main effects. This suggests that when both forms of discounts are used in the same offer, there is only a slight lift in conversion over the level expected when just one of two discounts is included. Another noteworthy finding was that new versions of the graphic and message actually reduced conversion rates, all else constant.

```
Parameter Estimates

Effect              Estimate      Standard Error     DF        t Value      Pr > |t|
intercept           -3.4178       0.09771            31991     -34.98       <.0001
personal            -0.1255       0.08732            31991      -1.44       0.1507
graphic             -0.1566       0.06487            31991      -2.41       0.0158
message             -0.3240       0.09191            31991      -3.53       0.0004
copy                -0.0077       0.06482            31991      -0.12       0.9055
shipping             0.2809       0.09542            31991       2.94       0.0032
incentive            0.2653       0.09578            31991       2.77       0.0056
shipping*incentive  -0.2365       0.13050            31991      -1.81       0.0699
personal*message     0.1974       0.13030            31991       1.51       0.1299
```

Output 2. Parameter Estimates.

OBTAINING PREDICTIONS FROM THE TEST RESULTS

The main goal of the experiment was to identify the combination of factor levels that would yield the highest predicted conversion rate for the rollout campaign. To do this we refit the model in the previous section, but also obtained conversion rate predictions for all 64 combinations of our 6 factors. A brief explanation of the SAS code and the output is given below.

First, we created a macro (factgen) that produces a data set with all possible combinations of a specified number of two-level factors. The macro takes two parameters as input, the number of factors (fac=6) and the name of the output data set (outdat=full_factorial).

```
/****************************************************************************/
/* Macro factgen:                                                           */
/* Creates a design matrix for a factorial experiment with N 2-level factors. */
/* fac = Number of factors                                                  */
/* outdat = Name of the output data set                                     */
/****************************************************************************/

%macro factgen(fac=, outdat=);
  %do i=1 %to &fac;
     data d&i ;
     v&i = 0; output;
     v&i = 1; output;
     run;
  %end;
```

```
  proc sql;
  create table &outdat as
  select * from
  d1 %do i=2 %to &fac; , d&i %end; ;
  quit;
%mend;

%factgen(fac=6,outdat=full_factorial)
```

Next, the variable names in the 'full_factorial' data set are renamed to match the factor names in the experiment and the data set is appended to the 'email' data set to create a new data set 'email2'. This data set contains the 32,000 observations from 'email' and the 64 observations from 'full_factorial'. Note that the values of Conversion (the outcome variable) are missing for all observations in the data set 'full_factorial'.

```
data full_factorial;
set full_factorial;
rename
v1 = personal
v2 = graphic
v3 = message
v4 = copy
v5 = shipping
v6 = incentive;
run;

data email2;
set email full_factorial;
run;
```

We used PROC GLIMMIX to fit the model parameters and write the predicted conversion rates to an output data set 'predicted'. The code is shown below. PROC GLIMMIX ignores observations with missing values of the dependent variable during model fitting, but calculates predicted values for all observations in the input data set. We obtained the desired predicted conversion rates along with their upper and lower 95% prediction limits by adding an OUTPUT statement to the PROC GLIMMIX step shown in the previous section.

```
proc glimmix data=email2;
model conversion(event='1') = personal graphic message copy shipping incentive
shipping*incentive personal*message / dist=binary link=logit s;
output out=predicted pred(ilink)=pred lcl(ilink)=lcl ucl(ilink)=ucl;
run;
```

The final results are then sorted and the top 10 observations are printed.

```
proc sort data=predicted;
by descending pred;
run;

title "Predicted Conversion Rates for Full Factorial Design";
title2 "Top 10 Combinations";
proc print data=predicted(obs=10 where=(conversion=.)) noobs split='*';
var personal graphic message copy shipping incentive pred lcl ucl;
label personal='Personal'
      graphic='Graphic'
      message='Message'
      copy='Copy'
      shipping='Shipping'
      incentive='Incentive'
      pred='Predicted *Conversion*Rate'
      lcl='LCL_95'
      ucl='UCL_95';
format pred lcl ucl percent7.2;
run;
```

The predicted conversion rates for the 64 factor level combinations ranged from 1.97% to 4.28%. Output 3 shows the 10 factor level combinations with the highest predicted conversion rates. The combination with the highest conversion rate had Shipping and Incentive set to level 2 (coded as 1) and all other factors set to level 1 (coded as 0). This combination was selected as the winner for the rollout campaign. Because several of the combinations have very similar predicted conversion rates, an alternative to selecting a single winning combination is to select a few of the top combinations for further testing during the rollout campaign.

```
Predicted Conversion Rates for Full Factorial Design
Top 10 Combinations


                                                  Predicted
                                                  Conversion
Personal Graphic Message Copy Shipping Incentive  Rate      LCL_95   UCL_95

    0       0       0      0      1        1       4.28%     3.59%    5.08%
    0       0       0      1      1        1       4.25%     3.59%    5.02%
    0       0       0      0      1        0       4.16%     3.49%    4.96%
    0       0       0      1      1        0       4.13%     3.47%    4.92%
    0       0       0      0      0        1       4.10%     3.43%    4.90%
    0       0       0      1      0        1       4.07%     3.42%    4.84%
    1       0       0      0      1        1       3.79%     3.18%    4.52%
    1       0       0      1      1        1       3.76%     3.15%    4.49%
    1       0       0      0      1        0       3.69%     3.10%    4.39%
    0       1       0      0      1        1       3.68%     3.09%    4.38%
```

Output 3. Printout of the 10 Factor Level Combinations Having the Highest Predicted Conversion Rates.

CONCLUSION

We demonstrated how to design and analyze a simple fractional factorial marketing experiment with JMP and SAS[®] software using data from an email marketing campaign. We also showed how the test results were used to build a predictive model to guide marketing decisions. We hope this example will be valuable to other marketing statisticians who need to design and analyze results of their own marketing experiments and to practitioners in other fields who can use the same general approach to the design and analysis of fractional factorial tests.

REFERENCES

- Almquist, E., and Wyner, G. (2001), "Boost Your Marketing ROI with Experimental Design," *Harvard Business Review,* October 1, 2001.

- Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978), *Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building,* New York: John Wiley & Sons, Inc.

- Kuehl, R. O. (2000), *Design of Experiments: Statistical Principles of Research Design and Analysis* (2nd ed.), Pacific Grove, CA: Brooks/Cole.

- Ledolter, Johannes, and Swersey, Arthur J. (2007), *Testing 1-2-3: Experimental Design with Applications in Marketing and Service Operations*, Stanford, CA: Stanford University Press.

- SAS Institute Inc., (2009), *JMP[®] Design of Experiments*, Cary, NC: SAS Institute Inc.

- SAS Institute Inc., (2009), *SAS/STAT[®] User's Guide, Version 9*, Cary, NC: SAS Institute Inc.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Robert Krutsick
Enterprise: True Action Network
Address: 1075 First Avenue
City, State ZIP: King of Prussia, PA 19425
Work Phone: 610-491-4157
E-mail: krutsickr@trueaction.com