Class 13: Design of Marketing Experiments/ Multi Variable Testing

Professor Tae Jung Yoon (윤태중)
KAIST College of Business

Marketing Research & Analytics

Marketing Research & Analytics Course Structure

Getting Ready for Marketing Research and Analytics

- Marketing Research and Analytics Overview (Class 1)
- How to Tell Good From Bad Data Analytics (Class 2)
- Using Stata for Marketing Research and Analytics (Classes 2 & 3)
- Statistics Review (Class 4)

Understanding Customers and Markets

- Quantifying Customer Value (Class 1)
- Case Analysis: "Home Alarm, Inc.: Assessing Customer Lifetime Value" and Testing (Class 3)
- Measuring Customers' Willingness to Pay (Class 6)
- Valuation of Products: Conjoint Analysis (Classes 8 & 9)
- Market Segmentation: Cluster Analysis (Class 10)
- Survey, and Qualitative Research (Class 10)

Prospecting and Targeting the Right Customers

- Predicting Response with RFM analysis (Class 5)
- Case Analysis: "Tuango: RFM Analysis for Mobile App Push Messaging"; Lift and Gains (Class 6)
- Predicting Response with Logistic Regression (Class 7)
- Case Analysis: "BookBinders: Predicting Response with Logistic Regression" (Class 8)
- Predicting Response with Neural Networks (Class 9)
- Predicting Response with Decision Trees (Class 10)

Developing Customers

- Case Analysis: "Intuit: Quickbooks Upgrade" (Class 11)
- Next-Product-To-Buy Models: Learning From Purchases (Class 11)
- Recommendation Systems: Learning From Ratings (Class 12)

Retaining Customers

- Predicting Attrition (Class 12)

Selecting the Right Offers

- Design of Experiments / Multi Variable Testing (Class 13)
- Case Analysis: "Capital One: Information-Based Credit Card Design" (Class 14)

Limitations of Marketing Analytics

- When Marketing Analytics, CRM, and Databases Fail (Class 14)

Wrap-up

- Wrap-Up (Class 14)

There are few more fundamental questions in business than “does it work?”

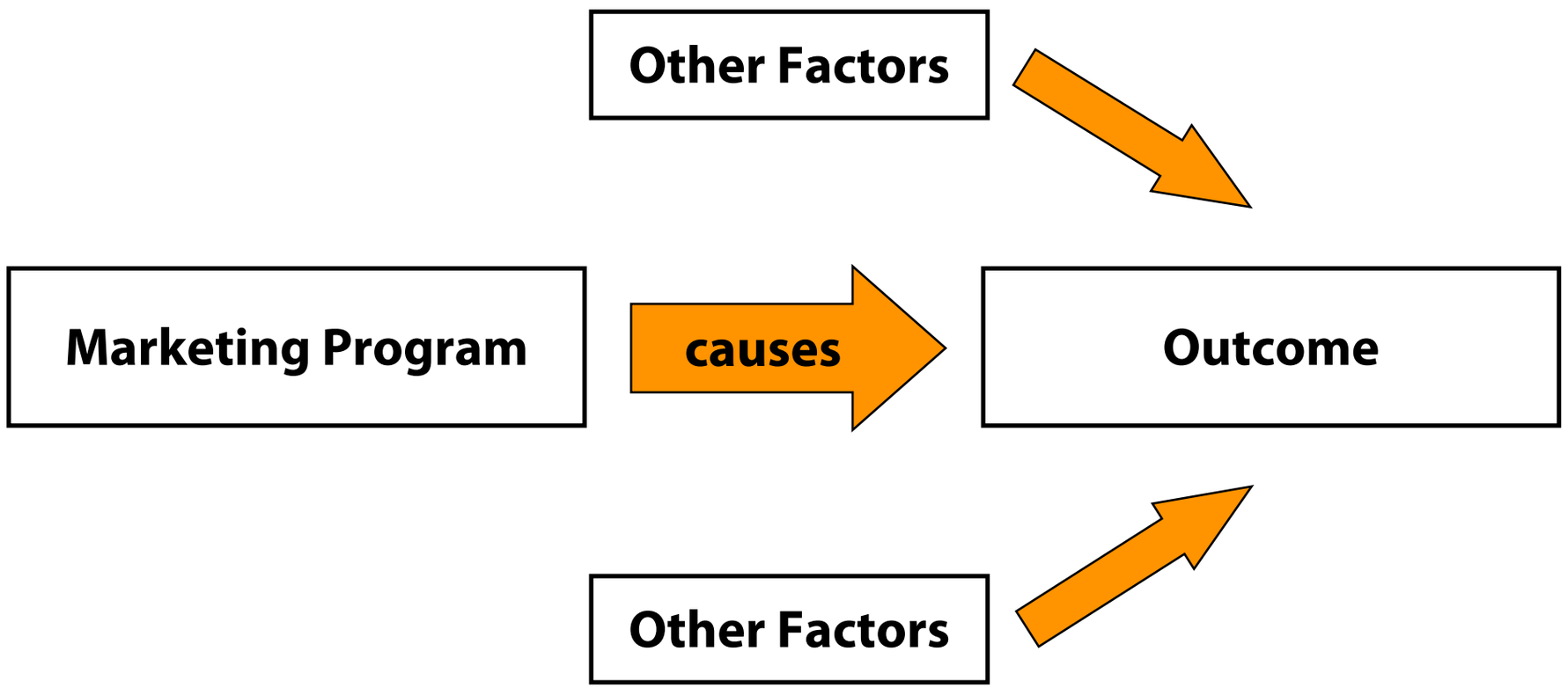
EXAMPLES OF MARKETING PROGRAMS

- Coupon
- Price promotion
- Advertising exposure
- Customer communication (sales call)
- Upgrade discount

- Purchase or not purchase
- Purchase quantity
- Retention
- Profitability

If you run a program and observe the outcome, can you conclude that the program caused the outcome?

CHALLENGE OF MEASURING EFFECT OF PROGRAM



How do we prove that an program worked?

WE NEED TO SHOW		THIS PROVES
If we run the program, the desired outcome occurs • Introduce promotion • Increase sales by 33%	→	The desired outcomes occurs when we run the program
If we do not run the program, the desired outcome does not occur • Don't introduce promotion • No increase in sales	→	The outcome was not caused by other factors

This is exactly what a test (experiment) does

The simplest experimental design is a two group random assignment

TWO GROUP RANDOM ASSIGNMENT

- Assign units (e.g. customers) randomly to receive or not receive treatment
- Treatment and control groups are not identical, because they consist of different individuals, but they are “probabilistically equivalent” meaning that there are not systematic differences between the groups in their characteristics or how they would respond to the program

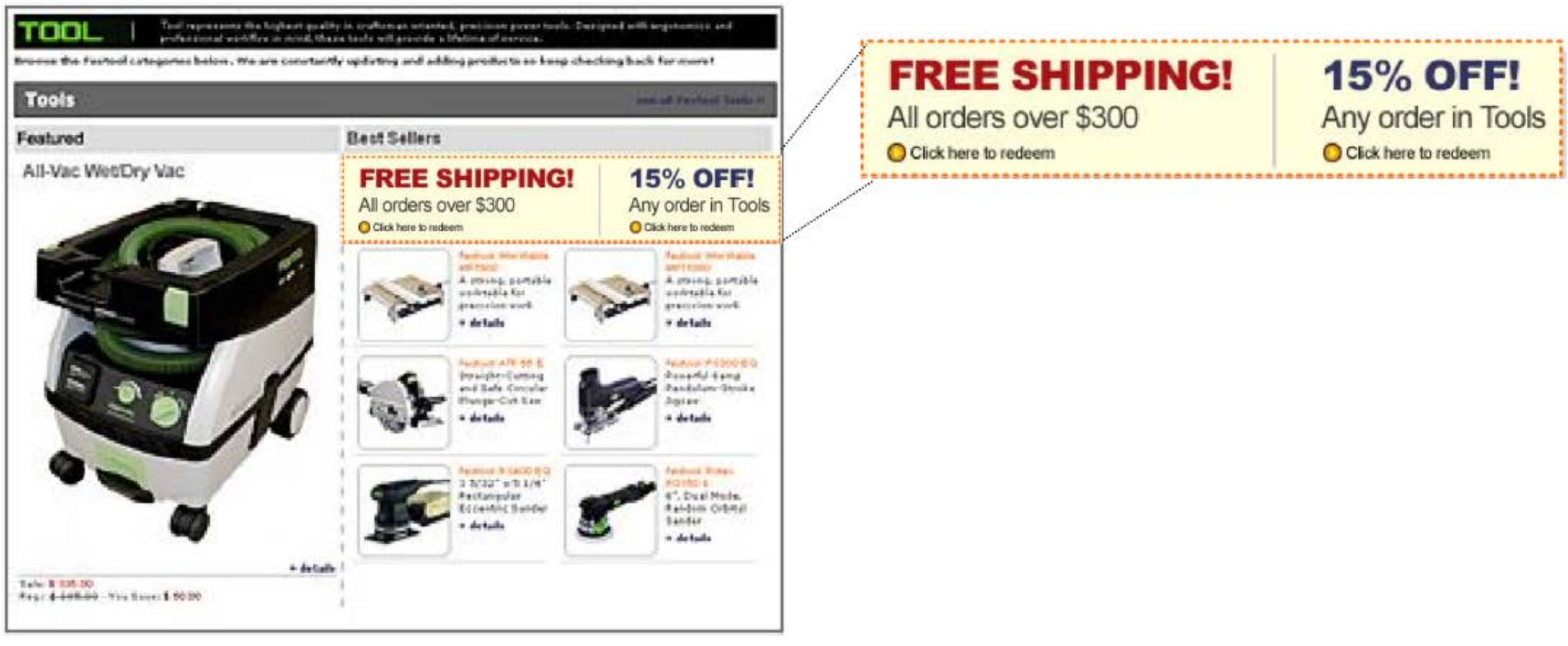
R O1
R X O2

- The difference in outcomes between treatment and control is the estimated effect of the program, **O2-O1**

Consider testing a change in the free shipping threshold for a website

EXAMPLE: FREE SHIPPING THRESHOLD

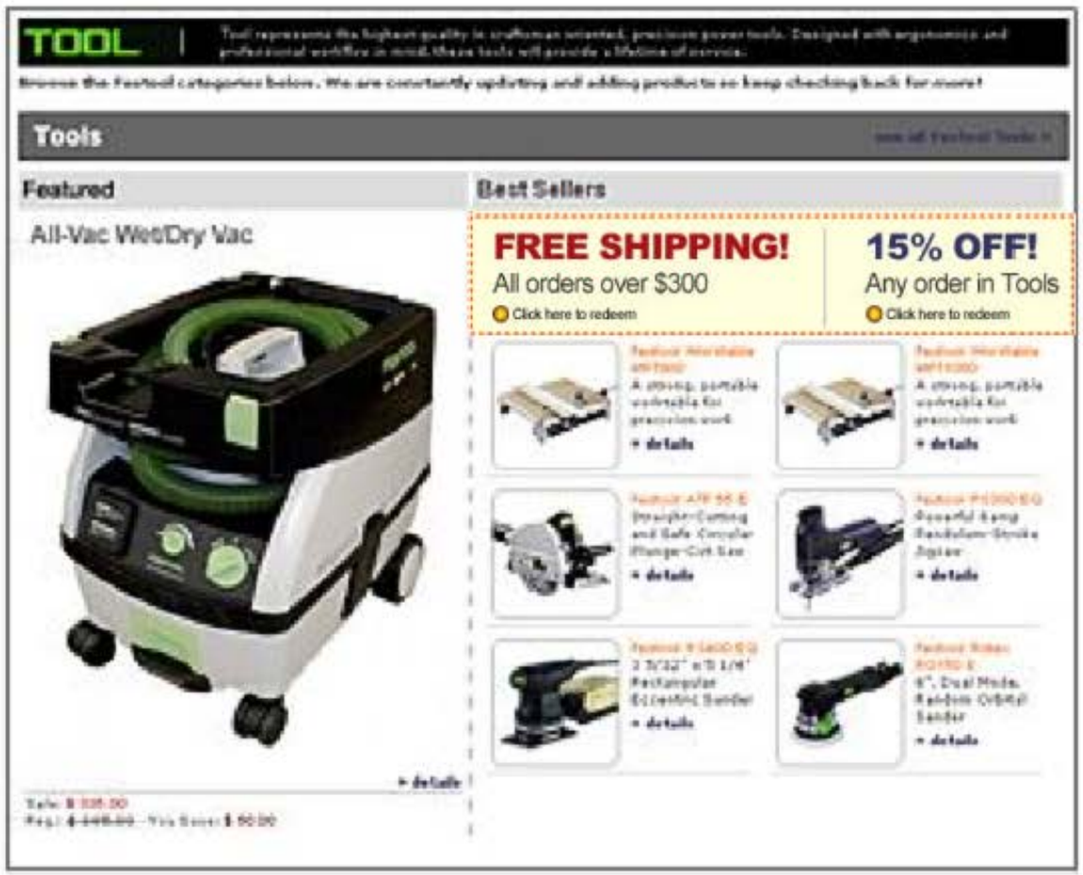
- Tools website, so far $300 threshold for free shipping



Consider testing a change in the free shipping threshold for a website

EXAMPLE: FREE SHIPPING THRESHOLD

- Tools website, so far $300 threshold for free shipping



We can now measure the effect of a lower shipping threshold

TWO GROUP RANDOM ASSIGNMENT

- Assign customers randomly to be exposed to
  - $300 (no treatment -- regular)
  - $200 (treatment)

R O1 --> $300 condition

R X O2 --> $200 condition

- The difference in outcomes between treatment and control is the estimated effect of the program, **O2-O1**

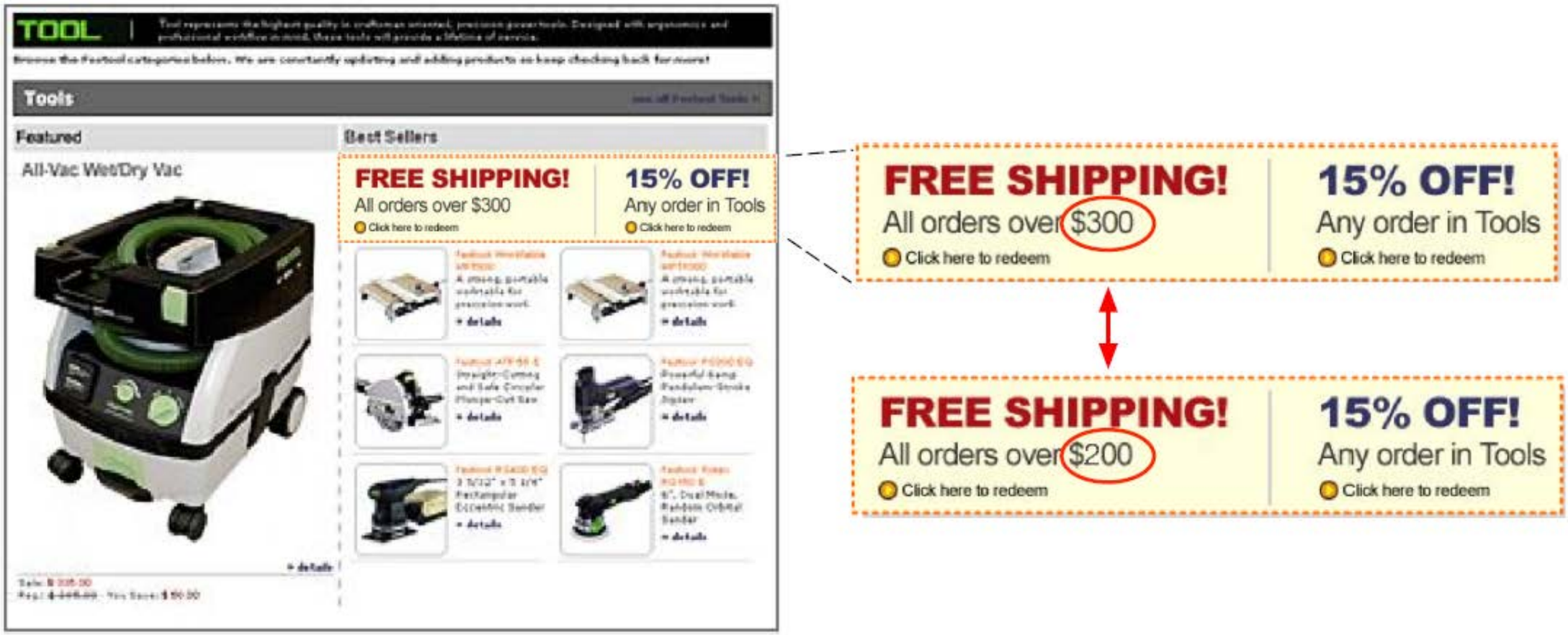
- Assume:
  - O1 is 500 sales for 10,000 exposures
  - O2 is 580 sales for 10,000 exposures

  --> Effect of a lower free shipping threshold is 80 sales

Often, we want to test more than one marketing idea

EXAMPLE: FREE SHIPPING THRESHOLD

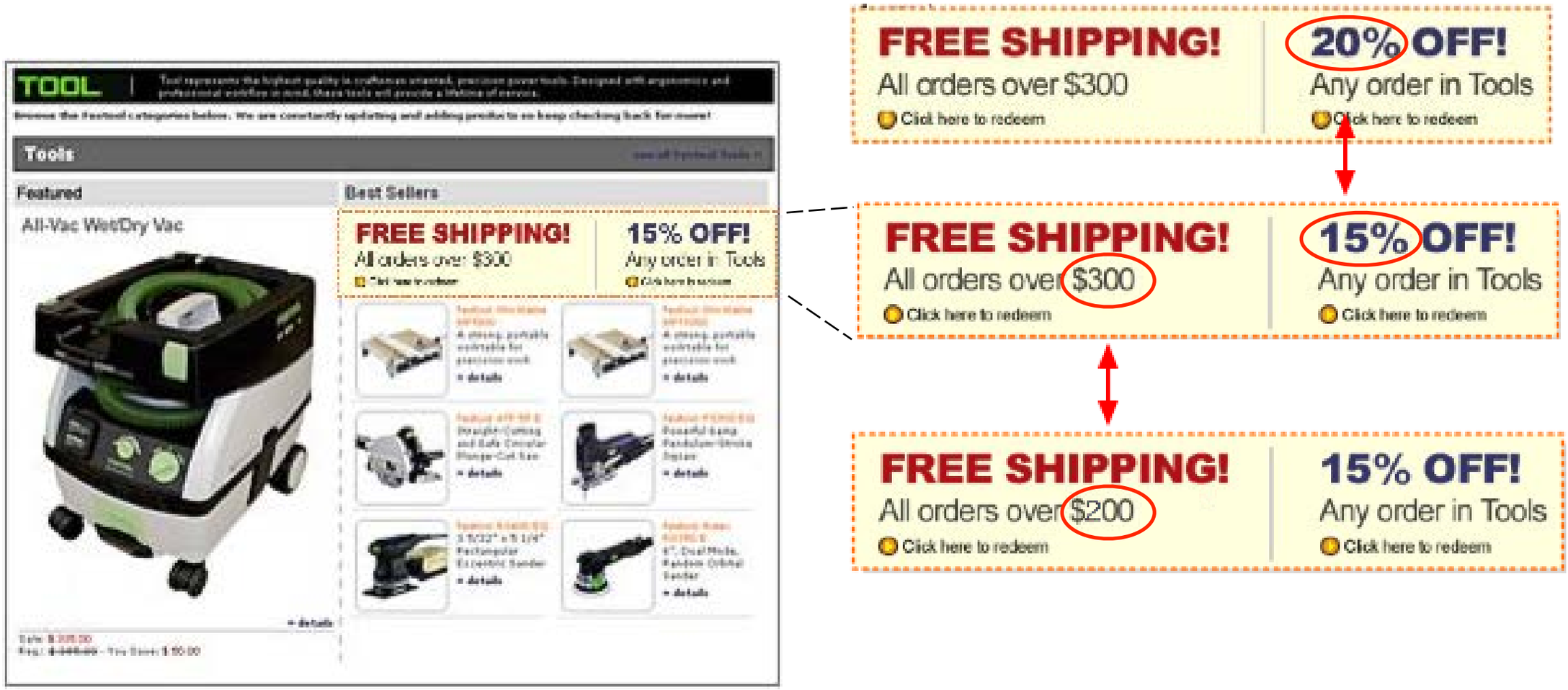
- Tools website, so far $300 threshold for free shipping



Often, we want to test more than one marketing idea

EXAMPLE: FREE SHIPPING THRESHOLD

- Tools website, so far $300 threshold for free shipping



We can run two separate experiments with a common control group

TWO GROUP RANDOM ASSIGNMENT

- Assign customers randomly to be exposed to
- $300, 15% (no treatment -- regular)
- $200, 15% (treatment 1)
- $300, 20% (treatment 2)

R O1 --> $300, 15% condition
R X1 O2 --> $200, 15% condition
R X2 O3 --> $300, 20% condition

- The difference in outcomes between treatments and control is the estimated effect of the program, **O2-O1** for treatment 1 and **O3-O1** for treatment 2

- Assume:
- O1 is 500 sales for 10,000 exposures
- O2 is 580 sales for 10,000 exposures
- O3 is 560 sales for 10,000 exposures

Effect of a lower free shipping threshold is **80 more sales**

Effect of a higher discount is **60 more sales**

A factorial design is a randomized experiment on multiple program features

FACTORIAL DESIGN

- Frequently you want to know not only
  - **whether** a program has an effect or not or **how big** the effect is,
  - but what **combination of features** will make the program most effective

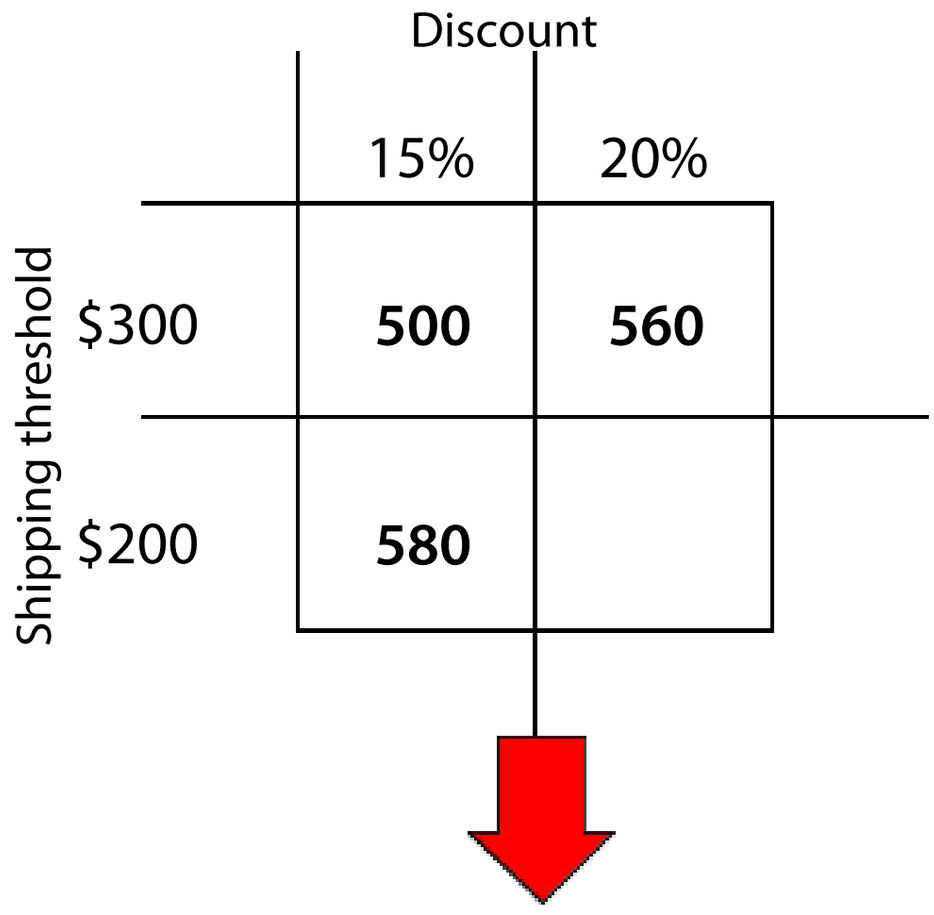
Experimental design:

- Factor 1: Free shipping threshold, (Xi)
- Factor 2: Tools discount (Xj)
  => "2x2 factorial design"
- Send one offers for each combination of factors to a sample of potential new customers
- Measure sales from first 10,000 exposures

R X11 O11
R X12 O12
R X21 O21
R X22 O22

The factorial design first simply replicates the two independent experiments

SALES (FOR 10,000 EXPOSURES) BY EXPERIMENTAL CONDITION



- Decrease the **shipping threshold** increase sales by **80**
- Increasing the **tools discount** increases sales **60**

The factorial design adds to simple experiments by also estimating whether there is an interaction effect

SALES (FOR 10,000 EXPOSURES) BY EXPERIMENTAL CONDITION

Shipping threshold	Discount 15%	Discount 20%
$300	**500**	**560**
$200	**580**	**680**

The factorial design adds to simple experiments by also estimating whether there is an interaction effect

SALES (FOR 10,000 EXPOSURES) BY EXPERIMENTAL CONDITION

Shipping threshold	Discount 15%	Discount 20%
$300	**500**	**560**
$200	**580**	**680**

- Decrease the **shipping threshold** increase sales by **80**
- Increasing the **tools discount** increases sales **60**
- Decreasing the **shipping threshold AND** increasing the **tools discount** increases sales by an additional **40** units

Full factorial designs can lead to too many "experimental conditions"

EXAMPLE OF MORE COMPLEX FACTORIAL DESIGN

Test an offer to attract Harrah's customer to the Las Vegas Property

Variations in the offer

- Factor 1: Number of free nights (1,2, or 3)
- Factor 2: Free chips ($0, $50, $100)
- Factor 3: Free show (no, yes)
- Factor 4: Expiration date (3 months, 6 months, 12 months)
- Factor 5: Follow-up phone call
- Factor 6: Follow-up e-mail

Objectives

- Maximize customer profitability over next year

3x3x2x3x2x2 factorial design = 216 different offers

Consider the following example to illustrate how to deal with too many experimental conditions

EXAMPLE: TOOLS WEBSITE CONTINUED

- Factor 1: Free shipping threshold ($200, $300)
- Factor 2: Discount (15%, 20%)
- Factor 3: Coupon code
  (Manual, automatic)

What is the sales effect of a lower shipping threshold?

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

What is the sales effect of a lower shipping threshold?

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

Sales effect of free shipping threshold

- With $300, sales are
  (500+520+560+580)/4 = 540

- With $200, sales are
  (580+600+640+660)/4 = 620

=> 620-540 = **80 incremental sales**

What is the sales effect of a higher discount?

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

Sales effect of an increased discount

- With 15%, sales are
  (500+520+580+600)/4 = 550

- With 20%, sales are
  (560+580+640+660)/4 = 610

=> 610-550 = **60 incremental sales**

What is the sales effect of automated coupon entry?

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

Sales effect of automated coupon entry

- With manual entry, sales are
  (500+560+580+640)/4 = 570

- With automated entry, sales are
  (520+580+600+660)/4 = 590

=> 590-570 = **20 incremental sales**

A partial factorial design uses fewer cells to calculate averages

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

PARTIAL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**

Idea:

- Cuts down the the number of offers and groups of customers that must be compared to estimate main effects.

We can still calculate the effect of a lower free shipping threshold

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

Sales effect of free shipping threshold

- With $300, sales are
  (500+520+560+580)/4 = 540
- With $200, sales are
  (580+600+640+660)/4 = 620

=> 620-540 = **80 incremental sales**

PARTIAL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**

Sales effect of free shipping threshold

- With $300, sales are
  (500+580)/2 = 540
- With $200, sales are
  (600+640)/2 = 620

=> 620-540 = **80 incremental sales**

The partial factorial approach generates the same result

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

Sales effect of an increased discount

- With 15%, sales are
  (500+520+580+600)/4 = 550
- With 20%, sales are
  (560+580+640+660)/4 = 610

=> 610-550 = **60 incremental sales**

PARTIAL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**

Sales effect of an increased discount

- With 15%, sales are
  (500+600)/2 = 550
- With 20%, sales are
  (580+640)/2 = 610

=> 610-550 = **60 incremental sales**

The partial factorial approach generates the same result

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

Sales effect of automated coupon entry

- With manual entry, sales are
  (500+560+580+640)/4 = 570
- With automated entry, sales are
  (520+580+600+660)/4 = 590

=> 590-570 = **20 incremental sales**

PARTIAL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**

Sales effect of automated coupon entry

- With manual entry, sales are
  (500+640)/2 = 570
- With automated entry, sales are
  (580+600)/2 = 590

=> 590-570 = **20 incremental sales**

Consider the same effect but with slightly different sales results

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**620**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**700**

PARTIAL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	20%	Automatic	**620**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**

Sales effect of free shipping threshold

- With $300, sales are
  (500+520+560+620)/4 = 550

- With $200, sales are
  (580+600+640+700)/4 = 630

=> 630-550 = **80 incremental sales**

Consider the same effect but with slightly different sales results

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**620**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**700**

Sales effect of free shipping threshold

- With $300, sales are
  (500+520+560+620)/4 = 550
- With $200, sales are
  (580+600+640+700)/4 = 630

=> 630-550 = **80 incremental sales**

PARTIAL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	20%	Automatic	**620**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**

Sales effect of free shipping threshold

- With $300, sales are (500+620)/2 = 560
- With $200, sales are (600+640)/2 = 620

=> 620-560 = **60 incremental sales**

Consider the same effect but with slightly different sales results

FULL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**620**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**700**

Sales effect of free shipping threshold

- With $300, sales are
  (500+520+560+620)/4 = 550
- With $200, sales are
  (580+600+640+700)/4 = 630

=> 630-550 = **80 incremental sales**

PARTIAL FACTORIAL DESIGN

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 200.00	15%	Manual	**580**
$ 200.00	20%	Automatic	**700**

Sales effect of free shipping threshold

- With $300, sales are (500+620)/2 = 560
- With $200, sales are (600+640)/2 = 620

=> 620-560 = **60 incremental sales**

Sales effect of free shipping threshold

- With $300, sales are (520+560)/2 = 540
- With $200, sales are (580+700)/2 = 640

=> 640-540 = **100 incremental sales**

Problem?

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**580**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**660**

Free Shipping	Discount	Coupon Entry	Sales
$ 300.00	15%	Manual	**500**
$ 300.00	15%	Automatic	**520**
$ 300.00	20%	Manual	**560**
$ 300.00	20%	Automatic	**620**
$ 200.00	15%	Manual	**580**
$ 200.00	15%	Automatic	**600**
$ 200.00	20%	Manual	**640**
$ 200.00	20%	Automatic	**700**

We have been measuring the effect of the shipping cost threshold assuming that there are no interaction effects

INTERACTION EFFECT PROBLEM

Problem for partial factorial approach

- The selection of which cases to consider will result in different estimates for the effect of a factor, depending on the values of the other variables that matter for interactions
- We cannot recover the size of the interaction effect because we do not observe the necessary combination of variables

TOOLS EXAMPLE

- Free shipping threshold effect ($200 vs. $300): **80 sales**
- Discount effect (20% vs. 15%): **60 sales**
- Coupon effect (Automatic vs. manual): **20 sales**
- Coupon with discount interaction effect (20% and automatic): **40 sales**

A smart partial factorial design involves input from marketers

EXAMPLE OF FACTORIAL DESIGN PROBLEM

Test an offer to attract Harrah's customer to the Las Vegas property

Variations in the offer

- Factor 1: Number of free nights (1, 2, or 3)
- Factor 2: Free chips (0, $50, $100)
- Factor 3: Free show (no, yes)
- Factor 4: Expiration date
- Factor 5: Follow-up phone call
- Factor 6: Follow-up e-mail

Blocking

- We can only mail 10 different conditions per business day
- Know effectiveness varies by day when mailing received

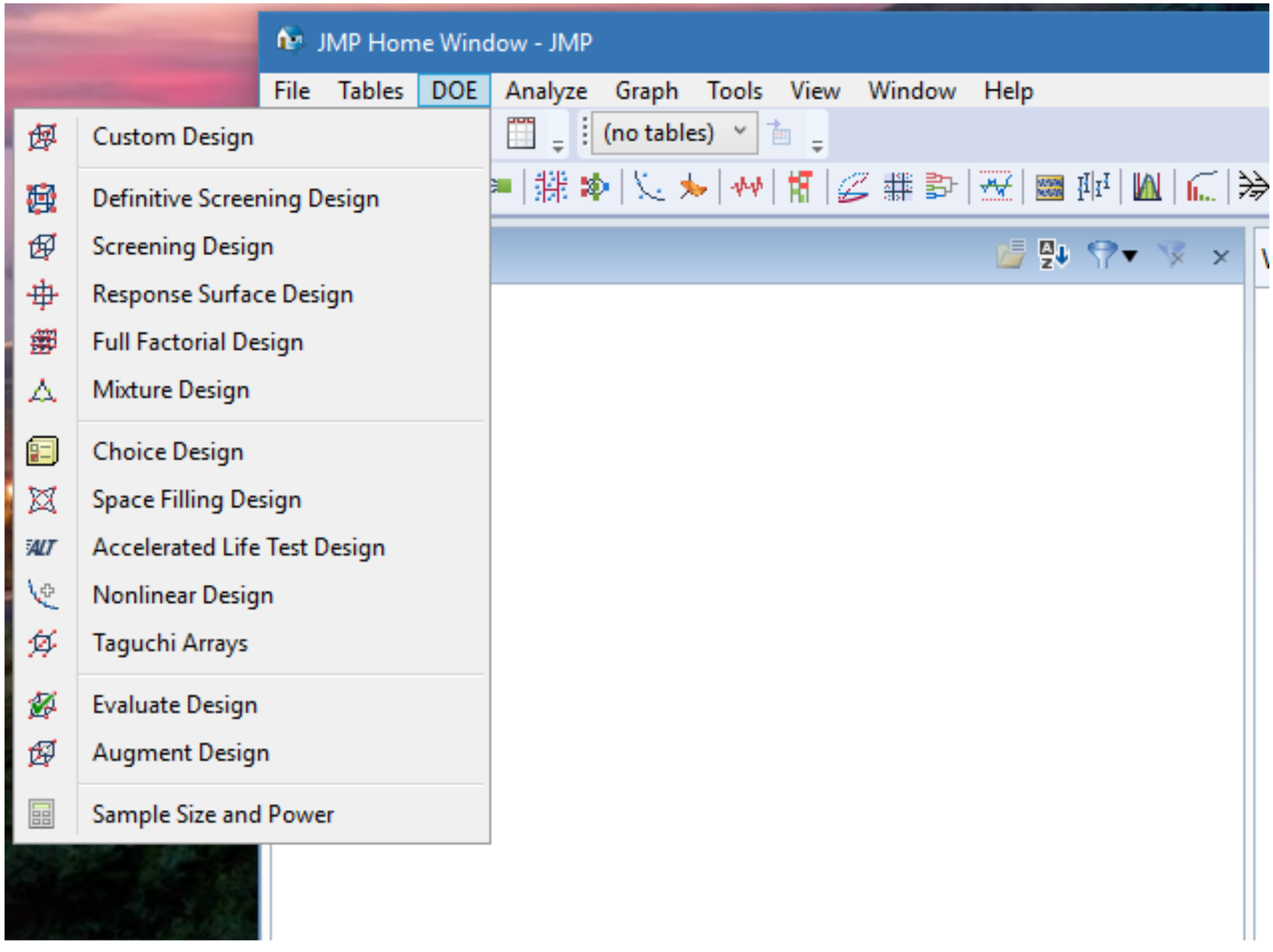
Objectives

- Maximize response rate
- Maximize customer profitability over next year

3x3x2x3x2x2 factorial design = 216 different offers

We can use JMP (SAS) to design the experiment

JMP DOE OPTIONS



There are a variety of tools for DOE

DOE TOOLS

JMP (SAS Institute)

Offermatica

- "Offermatica's software is a hosted testing and optimization platform that enables marketers to rapidly test offer content to maximize selling effectiveness"
- DOE tool for online offers with deployment

Others...

How do we analyze experiments in Stata?

ANALYSIS PROBLEM

Example: Assign customers randomly to be exposed to

- $300 (no treatment -- regular)
- $200 (treatment)

R O1 --> $300 condition

R X O2 --> $200 condition

- Result
  - O1 is 500 sales for 10,000 exposures
  - O2 is 580 sales for 10,000 exposures

  --> Effect of a lower free shipping threshold is 80 sales

Are consumers in the treatment ($200) condition *statistically significantly* more likely to purchase than in the control ($300) condition?

We have already covered the required methods

ANALYSIS IN STATA

	st200	sale
1	0	0
2	0	0
3	0	0
4	1	0
5	0	0
6	0	0
7	1	0
8	0	0
9	1	0
10	1	0
11	1	0
12	0	0
13	0	0
14	0	0
15	0	1
16	1	0
17	1	0
18	0	0
19	1	0
20	1	0
21	1	1

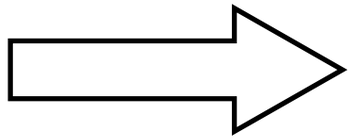
```
. tabulate sale st200, col chi2

           |         st200
      sale |         0          1 |     Total
-----------+----------------------+----------
         0 |     9,500      9,420 |    18,920
           |     95.00      94.20 |     94.60
-----------+----------------------+----------
         1 |       500        580 |     1,080
           |      5.00       5.80 |      5.40
-----------+----------------------+----------
     Total |    10,000     10,000 |    20,000
           |    100.00     100.00 |    100.00

          Pearson chi2(1) =   6.2642   Pr = 0.012
```

```
. logistic sale st200
Logistic regression                                     Number of obs   =      20000
                                                        LR chi2(1)      =       6.27
                                                        Prob > chi2     =     0.0123
Log likelihood = -4199.4386                             Pseudo R2       =     0.0007

------------------------------------------------------------------------------
        sale | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       st200 |   1.169851   .0733895     2.50   0.012     1.034502    1.322909
------------------------------------------------------------------------------
```

Can we avoid typing “0”s and “1”s 20,000 times in Stata?

ANALYSIS IN STATA

	st200	sale
1	0	0
2	0	0
3	0	0
4	1	0
5	0	0
6	0	0
7	1	0
8	0	0
9	1	0
10	1	0
11	1	0
12	0	0
13	0	0
14	0	0
15	0	1
16	1	0
17	1	0
18	0	0
19	1	0
20	1	0
21	1	1



	st200	sale	people
1	0	1	500
2	0	0	9500
3	1	1	580
4	1	0	9420

Instead of tabulate: `tabulate <sale> <experimental condition>, chi2`

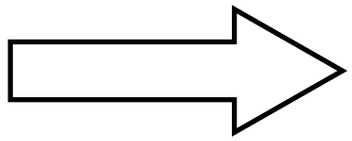
Use tabulate with frequency weights: `tabulate <sale> <experimental condition> [fw=people], chi2`

```
. tabulate sale st200 [fw=people], chi2

   (mean) |     (mean) st200
     sale |         0          1 |     Total
----------+----------------------+----------
        0 |     9,500      9,420 |    18,920
        1 |       500        580 |     1,080
----------+----------------------+----------
    Total |    10,000     10,000 |    20,000

          Pearson chi2(1) =   6.2642   Pr = 0.012
```

Can we avoid typing “0”s and “1”s 20,000 times in Stata?

ANALYSIS IN STATA

	st200	sale
1	0	0
2	0	0
3	0	0
4	1	0
5	0	0
6	0	0
7	1	0
8	0	0
9	1	0
10	1	0
11	1	0
12	0	0
13	0	0
14	0	0
15	0	1
16	1	0
17	1	0
18	0	0
19	1	0
20	1	0
21	1	1



	st200	sale	people
1	0	1	500
2	0	0	9500
3	1	1	580
4	1	0	9420

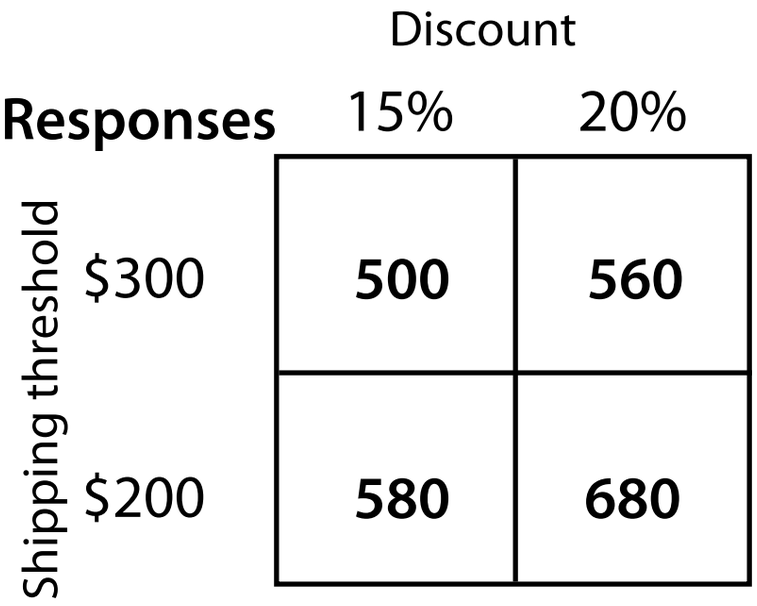
Instead of Logistic: `logistic <sale> <experimental condition>`

Use Logistic with frequency weights: `logistic <sale> <experimental condition> [fw=people]`

```
. logistic sale st200 [fw=people]

Logistic regression                                     Number of obs   =      20000
                                                        LR chi2(1)      =       6.27
                                                        Prob > chi2     =     0.0123
Log likelihood = -4199.4386                             Pseudo R2       =     0.0007

------------------------------------------------------------------------------
        sale | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       st200 |   1.169851   .0733895     2.50   0.012     1.034502    1.322909
------------------------------------------------------------------------------
```

This method can easily accommodate multiple experimental conditions

ANALYSIS IN STATA



	st200	dis20	sale	people
1	0	0	1	500
2	0	0	0	9500
3	0	1	1	560
4	0	1	0	9440
5	1	0	1	580
6	1	0	0	9420
7	1	1	1	680
8	1	1	0	9320

```
. generate st200_dis20=st200*dis20

. logistic sale st200 dis20 st200_dis20 [fw=people]

Logistic regression                                     Number of obs   =      40000
                                                        LR chi2(3)      =      30.39
                                                        Prob > chi2     =     0.0000
Log likelihood = -8841.9491                             Pseudo R2       =     0.0017

------------------------------------------------------------------------------
        sale | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       st200 |   1.169851   .0733895     2.50   0.012     1.034502    1.322909
       dis20 |   1.127119   .0712577     1.89   0.058     .9957624    1.275803
  st200_dis20 |   1.051347   .0904714     0.58   0.561     .8881738    1.244499
------------------------------------------------------------------------------
```

We can also use Stata to extrapolate from a partial factorial design or a partial design

PARTIAL FACTORIAL ANALYSIS IN STATA

Offer	Free Ship.	Discount	Coupon Entry	Sales
1	$ 300.00	15%	Manual	**500**
2	$ 300.00	15%	Automatic	**520**
3	$ 300.00	20%	Manual	**560**
4	$ 300.00	20%	Automatic	**620**
5	$ 200.00	15%	Manual	**580**
6	$ 200.00	15%	Automatic	**600**
7	$ 200.00	20%	Manual	**640**
8	$ 200.00	20%	Automatic	**700**

	offer	st200	dis20	ceaut	sale	people
1	1	0	0	0	1	500
2	1	0	0	0	0	9500
3	2	0	0	1	1	520
4	2	0	0	1	0	9480
5	3	0	1	0	1	560
6	3	0	1	0	0	9440
7	4	0	1	1	1	620
8	4	0	1	1	0	9380
9	5	1	0	0	1	580
10	5	1	0	0	0	9420
11	6	1	0	1	1	600
12	6	1	0	1	0	9400
13	7	1	1	0	1	640
14	7	1	1	0	0	9360
15	8	1	1	1	1	700
16	8	1	1	1	0	9300

We begin by adding a “partial” variable

PARTIAL FACTORIAL ANALYSIS IN STATA

Offer	Free Ship.	Discount	Coupon Entry	Sales
1	$ 300.00	15%	Manual	**500**
2	$ 300.00	15%	Automatic	**520**
3	$ 300.00	20%	Manual	**560**
4	$ 300.00	20%	Automatic	**620**
5	$ 200.00	15%	Manual	**580**
6	$ 200.00	15%	Automatic	**600**
7	$ 200.00	20%	Manual	**640**
8	$ 200.00	20%	Automatic	**700**

Offer	Free Ship.	Discount	Coupon Entry	Sales
1	$ 300.00	15%	Manual	**500**
4	$ 300.00	20%	Automatic	**620**
6	$ 200.00	15%	Automatic	**600**
7	$ 200.00	20%	Manual	**640**

	offer	partial	st200	dis20	ceaut	sale	people
1	1	1	0	0	0	1	500
2	1	1	0	0	0	0	9500
3	2	0	0	0	1	1	.
4	2	0	0	0	1	0	.
5	3	0	0	1	0	1	.
6	3	0	0	1	0	0	.
7	4	1	0	1	1	1	620
8	4	1	0	1	1	0	9380
9	5	0	1	0	0	1	.
10	5	0	1	0	0	0	.
11	6	1	1	0	1	1	600
12	6	1	1	0	1	0	9400
13	7	1	1	1	0	1	640
14	7	1	1	1	0	0	9360
15	8	0	1	1	1	1	.
16	8	0	1	1	1	0	.

We run our logistic regression on the tested cases and then predict purchase probabilities for the remaining cases

PARTIAL FACTORIAL ANALYSIS IN STATA

```
. logistic sale st200 dis20 ceaut [fw=people] if partial==1

Logistic regression                                     Number of obs   =      40000
                                                        LR chi2(3)      =      21.51
                                                        Prob > chi2     =     0.0001
Log likelihood = -8957.5294                             Pseudo R2       =     0.0012

------------------------------------------------------------------------------
        sale | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       st200 |   1.120073   .0477418     2.66   0.008     1.030302    1.217664
       dis20 |   1.159876   .0494384     3.48   0.001     1.066916    1.260936
       ceaut |   1.082757   .0461512     1.87   0.062     .9959772    1.177097
       _cons |   .0526316   .0024149   -64.17   0.000      .048105     .057584
------------------------------------------------------------------------------

. predict prob_partial

. gen numsales= prob_partial*10000
```

Stata Demo

Marketing Research & Analytics Course Structure

Getting Ready for Marketing Research and Analytics

- Marketing Research and Analytics Overview (Class 1)
- How to Tell Good From Bad Data Analytics (Class 2)
- Using Stata for Marketing Research and Analytics (Classes 2 & 3)
- Statistics Review (Class 4)

Understanding Customers and Markets

- Quantifying Customer Value (Class 1)
- Case Analysis: "Home Alarm, Inc.: Assessing Customer Lifetime Value" and Testing (Class 3)
- Measuring Customers' Willingness to Pay (Class 6)
- Valuation of Products: Conjoint Analysis (Classes 8 & 9)
- Market Segmentation: Cluster Analysis (Class 10)
- Survey, and Qualitative Research (Class 10)

Prospecting and Targeting the Right Customers

- Predicting Response with RFM analysis (Class 5)
- Case Analysis: "Tuango: RFM Analysis for Mobile App Push Messaging"; Lift and Gains (Class 6)
- Predicting Response with Logistic Regression (Class 7)
- Case Analysis: "BookBinders: Predicting Response with Logistic Regression" (Class 8)
- Predicting Response with Neural Networks (Class 9)
- Predicting Response with Decision Trees (Class 10)

Developing Customers

- Case Analysis: "Intuit: Quickbooks Upgrade" (Class 11)
- Next-Product-To-Buy Models: Learning From Purchases (Class 11)
- Recommendation Systems: Learning From Ratings (Class 12)

Retaining Customers

- Predicting Attrition (Class 12)

Selecting the Right Offers

- Design of Experiments / Multi Variable Testing (Class 13)
- Case Analysis: "Capital One: Information-Based Credit Card Design" (Class 14)

Limitations of Marketing Analytics

- When Marketing Analytics, CRM, and Databases Fail (Class 14)

Wrap-up

- Wrap-Up (Class 14)

Next Time: Capital One Simulation

- Instructions in case and syllabus