

**data mining conference**

**M 2008**

data mining conference

M 2008

## Modeling & optimization marketing campaigns



Viterbo Berberena, Ph.D.  
Anahuac University (Mexico City)



Jaime Paredes, Computer Science Engineer  
Santander Bank in Mexico, Payment Systems

# Presentation overview

- Data Collection
- Research methodology
  - Marketing Testing
  - Marketing Optimization
- Conclusions

# Presentation overview

- Data Collection
- Research methodology
  - Marketing Testing
  - Marketing Optimization
- Conclusions

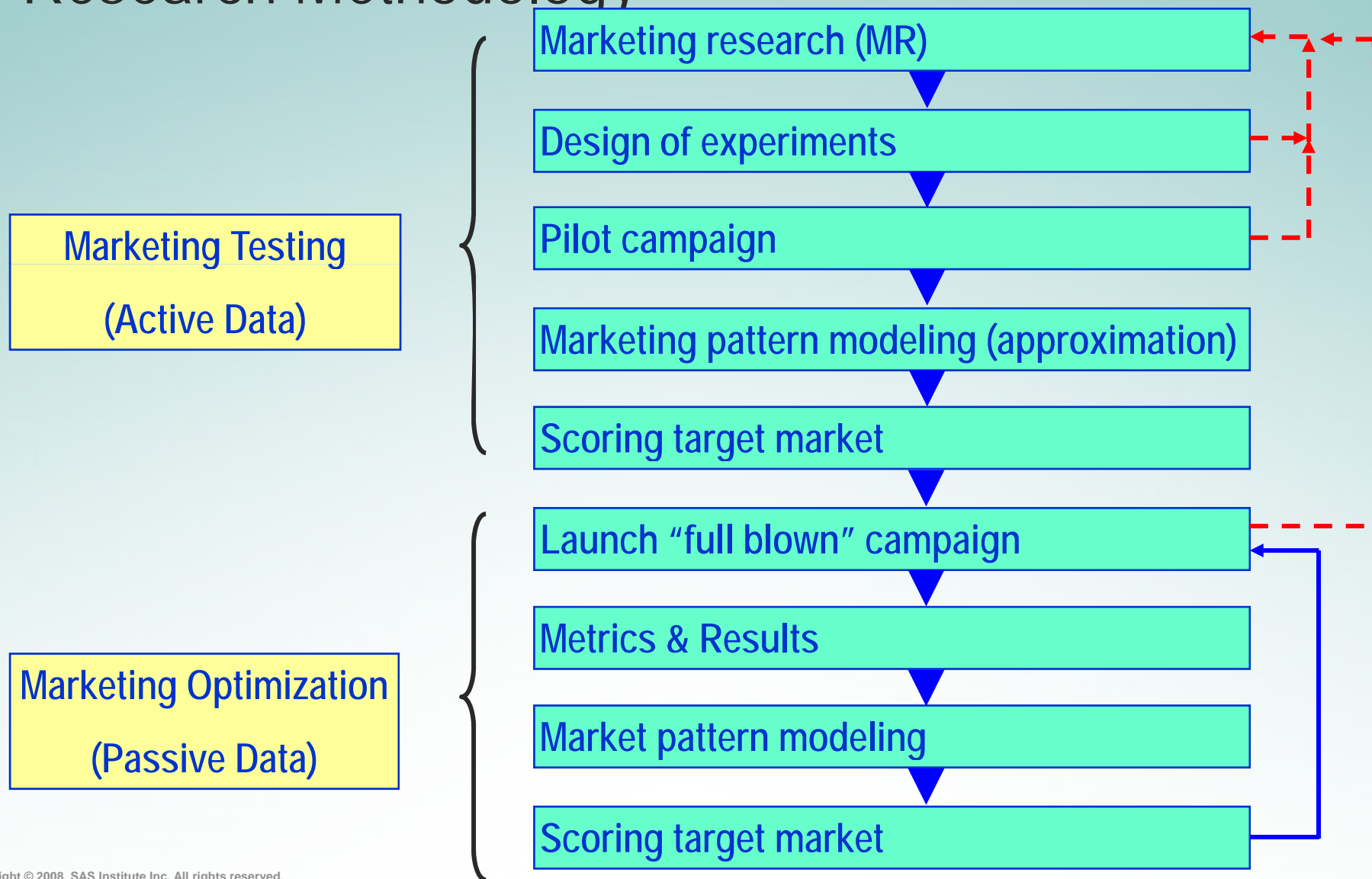
# Data Collection

- We need quantitative data (facts not symptoms) from customers !!
- Why??
  - To develop marketing models to support for better decision making (profitable decisions)
- How??
  - Depending the marketing approach, but also where data is located:
    - **Passive data**: stored in data warehouse, datamarts, others possible sources, etc.
    - **Active data**: Extracted quantitatively from customer behavior testing

# Presentation overview

- Data Collection
- Research methodology
  - Marketing Testing
  - Marketing Optimization
- Conclusions

# Research Methodology



# Presentation overview

- Data Collection
- Research methodology
  - Marketing Testing
  - Marketing Optimization
- Conclusions



# Marketing Research (MR)

- **One classical definition:**

- MR is a systematic and objective identification, collection, analysis, dissemination, and use of information for the purpose of improving decision making related to the identification and solution of problems (and opportunities) in marketing

Source: Marketing Research, An applied orientation, by Naresh K. Malhotra, Prentice Hall

- **American Marketing Association new definition:**

- MR is the function that links the consumer, customer, and public to the marketer through information to identify and define marketing opportunities and problems...

## Marketing Research (MR), cont'd

- To provide qualitative understanding
  - Discovery of ideas & insight
  - Get some marketing clues about customers:
    - Opinions, intentions, attitudes, awareness, motivations, perceptions, preferences, feelings, etc.
  - All them are just symptoms not facts!!
- From MR subjective discoveries
  - Formulate marketing hypothesis to be tested
  - Testing with quantitative approach: DOE

# Marketing Research (MR), cont'd

- Methods:
  - Qualitative research
    - Focus group (interviews) pre or post to campaign
    - Pre or post campaign depth interviews
  - Quantitative research
    - Pre campaign survey
    - Post campaign survey

# Design of Experiments (DOE)

- **Classical approach**: One factor at a time
  - Slow and expensive to perform research
  - Change one factor at a time and fix the rest
  - It fails to consider possible variable interaction, but also, without statistical significance to associate cause and effect
  - Besides these disadvantages it is widely used in practice
  
- **Statistical design of experiments**: DOE
  - Speed, low cost, highly reliable
  - Objective (quantitative) approach to analysis
  - The three basic principles:
    - Randomization
    - Replication
    - Blocking

## Design of Experiments (DOE) , cont´d

- DOE evaluates simultaneously several factors:
  - Marketing drivers (like promotions, channels, etc.)
  - Each experiment = micro campaign
  - The experimental matrix is a combinatorial of all levels of the factors (independent variables)
  - Different schools of DOE:
    - R. A. Fisher the original approach (UK)
    - G.E.P Box and Wilson, Hunter, Montgomery and others (US)
    - V.V. Kafarov, Nalimov and others (Russia)
    - G. Taguchi (Japan)
    - A.J. Swersey from Yale University has a formal work and a Book with applied DOE in marketing

# Design of Experiments (DOE) , cont'd

**Full Factorial 3<sup>2</sup> Design Matrix**

Trial Number	Coded Variables		Natural Variables		Sample Size	Response Level %
	X <sub>1</sub>	X <sub>2</sub>	Driver	Segment		
1	+1	+1	Reminder	Segment 1	3,000	7.63%
2	0	+1	Promo 1	Segment 1	3,000	7.50%
3	-1	+1	Promo Plus	Segment 1	3,000	7.83%
4	+1	0	Reminder	Segment 2	3,000	4.10%
5	0	0	Promo 1	Segment 2	3,000	4.70%
6	-1	0	Promo Plus	Segment 2	3,000	4.07%
7	+1	-1	Reminder	Segment 3	3,000	2.67%
8	0	-1	Promo 1	Segment 3	3,000	2.47%
9	-1	-1	Promo Plus	Segment 3	3,000	2.07%
Control Group (Business as Usual Effect)					3,000	4.90%
<b>Total</b>					<b>30,000</b>	

- **Theoretical response surface**

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 + b_4x_1^2 + b_5x_2^2$$

## Design of Experiments (DOE) , cont´d

- If outcomes for the experiments fit a response surface, it is possible to establish:
  - Among proven factors, Which ones influence significantly the results of the campaign?
  - Which is the factor ranking?
  - Within each factor, Which is the hierarchical order of level's significances?
  - Which are the best marketing conditions for the launch of a efficient and effective pilot campaign.
  - But!! Even if a response surface is found, but from marketing perspective is usefulness, it is required to refine once again with another design of experiments



# Design of Experiments (DOE) , cont'd

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	0.00418778	0.00083756	108.72	0.0014
Error	3	0.00002311	0.00000770		
Corrected Total	8	0.00421089			

R-Square	Coeff Var	Root MSE	Y Mean
0.994512	5.795822	0.002776	0.047889

Parameter	Estimate	Standard Error	t Value	Pr >  t
Intercept	0.0441111111	0.00206878	21.32	0.0002
X1	0.0006666667	0.00113312	0.59	0.5977
X2	0.0260000000	0.00113312	22.95	0.0002
X1*X2	-.0020000000	0.00138778	-1.44	0.2452
X1*X1	-.0016666667	0.00196261	-0.85	0.4582
X2*X2	0.0073333333	0.00196261	3.74	0.0334

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 + b_4x_1^2 + b_5x_2^2$$

$$y = 0.04411 + 0.026x_2 + 0.00733x_2^2$$



## Design of Experiments (DOE) , cont´d

- If outcomes for the experiments do not fit to the response surface:
  - Marketing factors tested, no cause and effect in the campaign outcomes under the experiment conditions
  - Campaign response behavior is independent from selected marketing drivers for the experiments
  - Go back to marketing research and carry out one or both of the following actions:
    - Analyze why the experiment failed:
      - Conditions, target segments, timing, seasonality, ...
    - Test new drivers as experiment factors:
      - Promotion attractiveness, marketing creativity, distribution channels, ...

# Pilot Campaign

- **Pilot campaign target:** Supply data (facts, quantitative evidence ) to extract an approximation of the customer's response pattern
- Marketing conditions established from DOE (e.g. promo, channel, target segment)
- Pilot campaign size, ought to be bigger, aprox. 10 times DOE's micro-campaigns (experimental sample size)

## Marketing pattern modeling (approximation)

- From the outcomes of the marketing pilot campaign (level response 0s & 1s), a supervised model is trained.
- For each customer from pilot campaign a row vector is added. Each vector is integrated by:
  - Inputs: Business behavior variables
  - Target: Flagged binary response variable ( 0s & 1s)
- If a model can be trained, it represents an approximation of the actual customer behavior
- Lack of fit of the model, is due to the pilot campaign size, which is still small for extracting a stronger mathematical regularity (pattern)

## Scoring target market

- **Scoring is the last step of the first phase of this methodology (marketing testing – active data)**
- The response model trained allows the scoring of the target market (customer's segments, etc.)
- As a result of the scoring process, it is important to consider the probability of responsiveness from a cut-off and above (greater propensity to respond positively)

# Presentation overview

- Data Collection
- Research methodology
  - Marketing Testing
  - Marketing Optimization
- Conclusions

## Launch “full blown” campaign

- The first step of marketing optimization (methodology’s second phase)
- “Full blown” campaign is a real marketing campaign, is no longer an experiment, has all elements and conditions to target customers from behavioral and marketing perspectives.
- It’s size is at least 10 times greater than the pilot campaign.
- However, it is important to consider not to launch this campaign too big in size, because:
  - Pattern from propensity models is still an approximation, it has some degree of lack of fit
  - Increasing the size of the campaign should also consider a cost & profit approach

## Metrics & Results

- Group and individual Post-Prev analysis can be performed to analyze the response (ANOVAs, t-test, Wilcoxon-Mann Whitney test among others)
- The post-prev analysis at individual level is used to determine quantitatively “responders” and “non-responders”.
- If the discipline of using a control group for campaigns is considered, market inertia (business as usual effect) can be added or subtracted from the results to obtain the net effect of the campaign
- The expected values of the business variables can be also calculated to have them as references for subsequent campaigns of the same type.



# Market pattern modeling

- From “full blown” campaign results, a new supervise model is trained
- The new model captures a more accurate pattern of the actual customer behavior
- The increase of accuracy, comes from the size of the “full blown” campaign
- The goodness of fit is better because the size of the campaign allows the extraction of a stronger mathematical regularity.



## Scoring target market

- **Once again “scoring process” is the last step of the second phase of this methodology (marketing optimization – passive data)**
- More accurate scoring model than the first one extracted from DOE (approximated pattern)
- Greater effectiveness can be achieved because of better pattern recognition:
  - Better probabilities produce better customer targeting for subsequent “full blown” campaigns
- A virtuous cycle takes place: “Optimization”:
  - Campaign Launch → Metrics & Results → Pattern refinement → Scoring target market → Campaign Launch → ...

# Presentation overview

- Data Collection
- Research methodology
  - Marketing Testing
  - Marketing Optimization
- **Conclusions**

# Conclusions

- The methodology is still under test and development phase at Santander
- This approach has proven to be an effective way to address customers
- Our concepts must not be considered the only approach to understand how to address customers from a marketing perspective, but is a good one.

## About us:

- Viterbo Berberena, Ph.D., Anahuac University (Mexico City),  
[vberberena@anahuac.mx](mailto:vberberena@anahuac.mx)
- Jaime Paredes, Computer Science Engineer  
Santander Bank in Mexico  
[japaredes@santander.com.mx](mailto:japaredes@santander.com.mx)