# MARKETING ANALYTICS I

*Measurement:*
*Marketing Mix Models and Attribution Models*

# *Introduction*

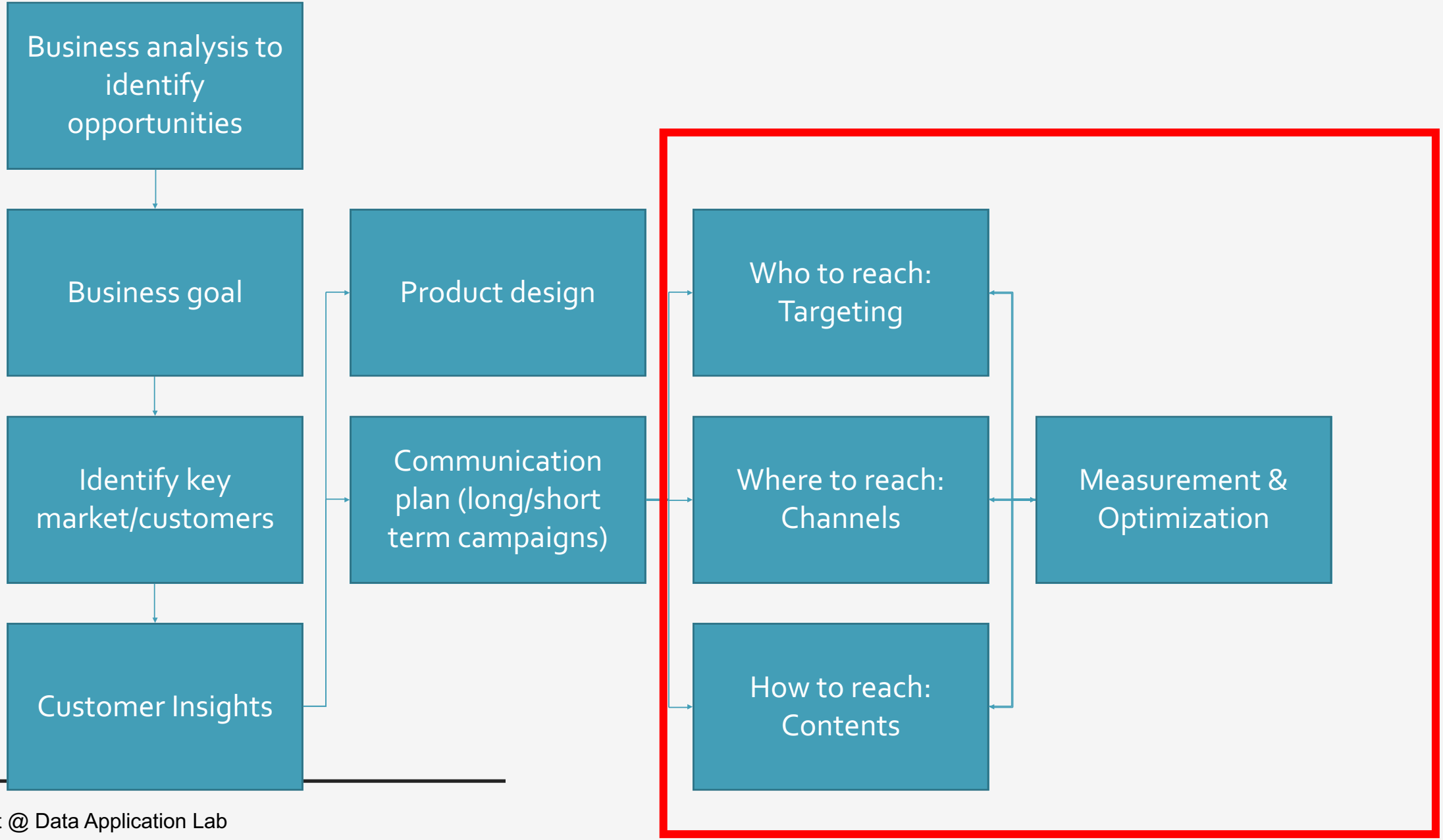
- Marketing analytics opportunities and potential career paths
- What does it take?
  - Business acumen
  - Technical skills
  - Industry knowledge
  - Communication and visualization

Copyright @ Data Application Lab

# Contents


- Marketing ecosystem and data application in marketing industry
- Frameworks
- Special Topic: Measurement and Optimization
  - Marketing mix models
  - Attribution models
  - Causal inference (A/B tests and forensic control)


Copyright @ Data Application Lab




Copyright @ Data Application Lab

| | |
|---|---|
| Business analysis to identify opportunities | SWOT, financial analysis, etc. |
| Business goal | |
| Identify key market/customers | Customer Segmentation, LTV analysis, etc. |
| Customer Insights | Quantitative & qualitative |

Copyright @ Data Application Lab

| | | |
|---|---|---|
| Business analysis to identify opportunities | SWOT, financial analysis, etc. | 1st party and 3rd party data (Panel, Industry trend report, retail data, etc.) |
| Business goal | | |
| Identify key market/customers | Customer Segmentation, LTV analysis, etc. | Quantitative data: demographics, spending behaviors, interests, etc. |
| Customer Insights | Quantitative & qualitative | Qualitative data (focus group, interviews, ethnographic research, etc.) |

Copyright @ Data Application Lab

| | | | |
|---|---|---|---|
| Business analysis to identify opportunities | SWOT, financial analysis, etc. | 1st party and 3rd party data (Panel, Industry trend report, retail data, etc.) | IRI, Neilson, Dunnhumby, Millward Brown, Worldpanel, etc. |
| Business goal | | | |
| Identify key market/customers | Customer Segmentation, LTV analysis, etc. | Quantitative data: demographics, spending behaviors, interests, etc. | KBM demo, DMP's |
| Customer Insights | Quantitative & qualitative | Qualitative data (focus group, in-depth interviews, ethnographic research, etc.) | |

Copyright @ Data Application Lab

| | |
|---|---|
| Who to reach: Targeting | Segmentation and LTV, DMP, DSP, RTB… |
| Where to reach: Channels | Measurement: MMM, attribution models, causal inference, A/B test |

Copyright @ Data Application Lab

| | | |
|---|---|---|
| Who to reach: Targeting | Segmentation and LTV, DMP, DSP, RTB… | Demographics, spending behaviors, interests, etc. |
| Where to reach: Channels | Measurement: MMM, attribution models, causal inference, A/B test | Transaction and impression, demo and other data if possible |

Copyright @ Data Application Lab

| | | | |
|---|---|---|---|
| Who to reach: Targeting | Segmentation and LTV, DMP, DSP, RTB… | Demographics, spending behaviors, interests, etc. | Starcom, Publicis, BlueKai |
| Where to reach: Channels | Measurement: MMM, attribution models, causal inference, A/B test | Transaction and impression, demo and other data if possible | Google Analytics, Oracle (Datalogix, MOAT), Adobe |

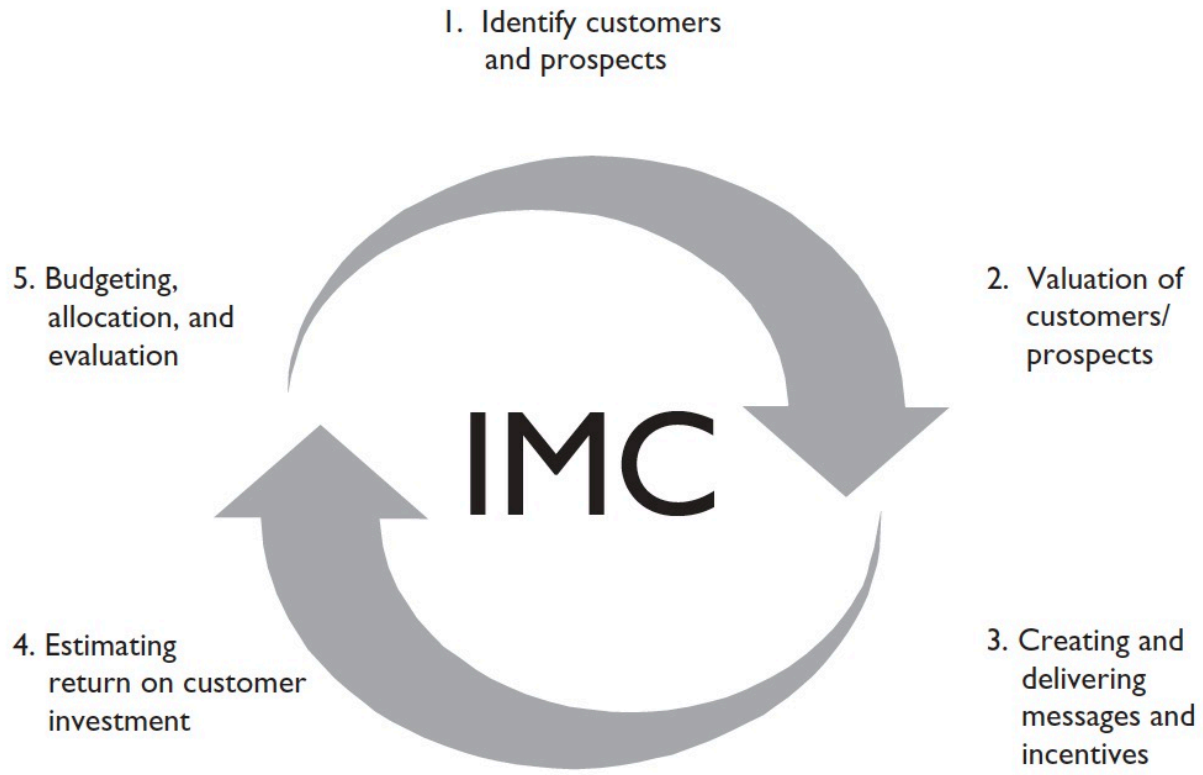
Service

OmnicomMediaGroup WPP OGILVY & MATHER MEDIACOM

Copyright @ Data Application Lab

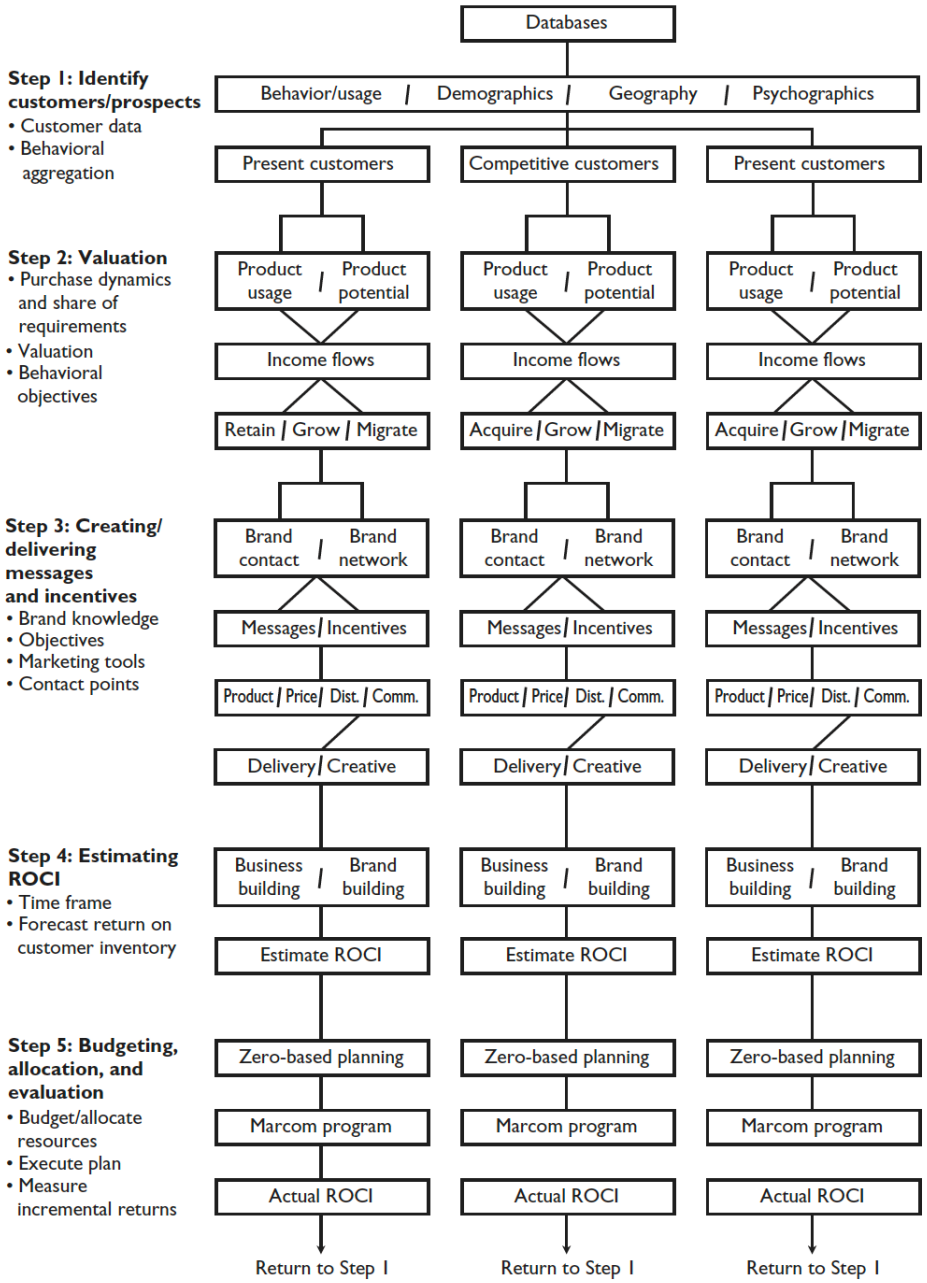
# *Integrated Marketing Communication*

**Exhibit 3.6 The Five-Step IMC Planning Process**




Copyright @ Data Application Lab

## Exhibit 3.7 Details of the Five-Step IMC Planning Process



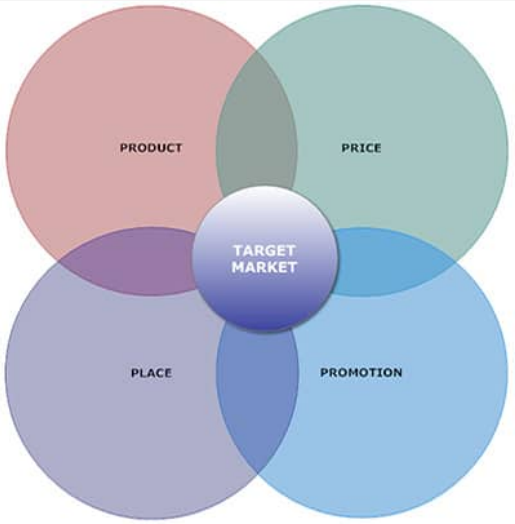

Note: Comm. = communication; Dist. = distribution.

Copyright @ Data Application Lab

# *MARKETING MIX MODELING*

Copyright @ Data Application Lab

# *4P's vs. 7P's*



Copyright @ Data Application Lab

# *MMM's answer the following questions:*

- How much media is enough?
- Which medium is most effective?
- What is the best media environment to use?
- When are ads worn out?

Copyright @ Data Application Lab

# *Marketing Mix Modeling*

Marketing mix modeling (MMM) is **statistical analysis** such as multivariate regressions

on **sales and marketing time series data**

to estimate the **impact of various marketing tactics** (marketing mix) on sales

and then **forecast** the impact of future sets of tactics.

Copyright @ Data Application Lab

# *Key metric: ROI*

- **Marketing Return on Investment** (MROI) will be a key metric to look at during your Marketing Mix Modeling project.

Copyright @ Data Application Lab

# *Step 1: Model Setup*

- Data aggregation
  - Weekly level being the most common
- Predicting variables
  - Radio ad investment, TV GRP (#impression*rating), online #impressions, etc.
- Predicted variable
  - Sales = **baseline sales** (sales when no promotions or offers are active) + **incremental sales** (due to marketing activities).

Copyright @ Data Application Lab

# *Step 2: Feature Selection*

- Brainstorm time: What other effects can impact weekly sales?

| Week | Sales | TV GRP | Online Impressions |
|---|---|---|---|
| 07/01/2017 | 30,503 | 0 | 22.82 M |
| 07/08/2017 | 27,037 | 0 | 0.00 M |
| 07/15/2017 | 30,646 | 0 | 0.00 M |
| 07/22/2017 | 40,887 | 0 | 0.00 M |
| 07/29/2017 | 48,947 | 0 | 0.00 M |
| 08/05/2017 | 37,910 | 100 | 127.65 M |
| 08/12/2017 | 40,436 | 93 | 125.34 M |
| 08/19/2017 | 49,343 | 95 | 150.62 M |
| 08/26/2017 | 32,371 | 90 | 206.28 M |
| 09/02/2017 | 28,665 | 12 | 595.09 M |
| 09/09/2017 | 29,079 | 17 | 284.73 M |
| 09/16/2017 | 22,794 | 11 | 46.09 M |
| 09/23/2017 | 26,607 | 0 | 13.62 M |
| 09/30/2017 | 21,153 | 0 | 36.95 M |

Copyright @ Data Application Lab

## SAMPLE MARKETING MIX MODEL SALES TACTIC

| DISTRIBUTION & PRICING | PRODUCT | PAID MEDIA | PROMOTIONS | EXTERNAL FACTORS |
|---|---|---|---|---|
| Distribution | Product Life Cycle | TV | Merchandising | Seasonality & Weather Patterns |
| Pricing | Product Changes | Magazines | Couponing | Competitive Factors |
| CRM & Offers | New Products | Newspaper | Public Relations | Macroeconomic Inputs |
| Channel Incentives | Segment Trends | Radio | Loyalty Program Activity | |
| Retail Format Changes | Product Recalls | Outdoor/Out-of-Home | Event Marketing | |
| | Quality Metrics | Online Media | Sponsorships | |
| | Awards | | Word-of-Mouth | |
| | Third-Party Reviews | | | |
| | Sampling | | | |
| | Inventory Levels | | | |
| | Sales Force Activity | | | |
| | Customer Satisfaction | | | |
| | Product Performance | | | |
| | Product Placement | | | |

Copyright @ Data Application Lab

| Week | Sales | Pricing | Distribution | Competition Discounts | Competition Online Impressions | TV GRP | Online Impressions | Promotions | Discounts |
|---|---|---|---|---|---|---|---|---|---|
| 07/01/2017 | 30,503 | $1,067 | 48 | 1.12% | 105.68 M | 0 | 22.82 M | 27 | 6.93% |
| 07/08/2017 | 27,037 | $1,068 | 47 | 4.33% | 0.00 M | 0 | 0.00 M | 5 | 8.55% |
| 07/15/2017 | 30,646 | $1,038 | 42 | 1.89% | 0.00 M | 0 | 0.00 M | 6 | 9.64% |
| 07/22/2017 | 40,887 | $954 | 35 | 1.10% | 0.00 M | 0 | 0.00 M | 6 | 13.75% |
| 07/29/2017 | 48,947 | $912 | 31 | 4.56% | 0.00 M | 0 | 0.00 M | 10 | 16.57% |
| 08/05/2017 | 37,910 | $1,010 | 38 | 3.64% | 66.62 M | 100 | 127.65 M | 12 | 11.15% |
| 08/12/2017 | 40,436 | $1,007 | 37 | 1.66% | 124.18 M | 93 | 125.34 M | 8 | 11.42% |
| 08/19/2017 | 49,343 | $994 | 33 | 2.50% | 96.87 M | 95 | 150.62 M | 10 | 12.90% |
| 08/26/2017 | 32,371 | $1,078 | 39 | 5.08% | 109.01 M | 90 | 206.28 M | 11 | 7.24% |
| 09/02/2017 | 28,665 | $1,060 | 40 | 1.10% | 115.16 M | 12 | 595.09 M | 2 | 7.20% |
| 09/09/2017 | 29,079 | $1,061 | 42 | 0.00% | 157.02 M | 17 | 284.73 M | 13 | 6.35% |
| 09/16/2017 | 22,794 | $1,098 | 41 | 0.15% | 145.53 M | 11 | 46.09 M | 7 | 5.74% |
| 09/23/2017 | 26,607 | $1,048 | 36 | 0.02% | 105.84 M | 0 | 13.62 M | 10 | 8.91% |
| 09/30/2017 | 21,153 | $1,100 | 39 | 0.00% | 118.05 M | 0 | 36.95 M | 11 | 5.99% |

Copyright @ Data Application Lab

# *Step 3: Model Selection*

## Approach 1: Multi-linear regression

Let $x_{it}$, i=1,…,K denote the K explanatory variables. These could include the price, measures of ad exposure, shelf placement, store location etc. To estimate the effect of each of the explanatory variables on the sales volume, we fit a regression model given by

$$y_t = \beta_0 + \sum_{i=1}^{K} \beta_i x_{it} + \varepsilon_t$$

where βi, i=0,… K denote the regression coefficients.

Copyright @ Data Application Lab

# *However, this is rarely the case…*

1. Exposure to advertising today has effects on consumers' purchasing decisions in the future (Carryover Effect)
2. The effect of ad exposure on purchasing decisions decays with time (Carryover Effect)
3. The effect of ad exposure on sales is subject to the law of diminishing returns (Shape Effect)
4. Changing the level of ad exposure brings about a relative change in sales volumes (Shape Effect)

Copyright @ Data Application Lab

# *Dealing with Carryover Effects*

- Approach 1: lag the effect for a certain period of time
- Approach 2: Create Adstock variable
  - Assuming a certain proportion (lambda) of the impact from one period can carry over to the next period:

$$y_t = \beta_0 + \beta_1 s_t + \varepsilon_t$$
$$s_t = x_t + \lambda s_{t-1}$$

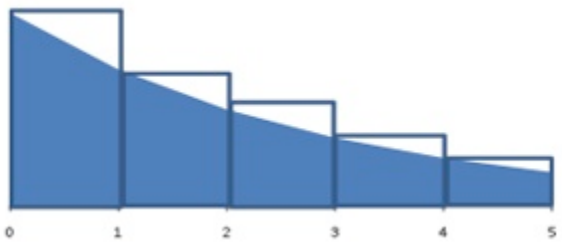
$$y_t = \beta_0 + \beta_1 x_t + \lambda \beta_1 s_{t-1} + \varepsilon_t$$

  - Spreadsheet in Excel

Copyright @ Data Application Lab

# *Using "half-life" to create Adstock Decay*

Week one = 120
Week two = 95
Week three = 76
Week four = 60
Week five = 48
Week six = 38




| | | |
|---|---|---|
| TV | 2-6 Week Half Life | Often awareness focused with delayed call to action |
| Radio | 1-5 Week Half Life | Varies substantially based on placement |
| Newspaper | 2-3 Week Half Life | Often Lower Funnel Focused |
| Print Magazines | 4-8 Week Half Life | Impact often delayed an elongated over life of periodical |
| Online (upper funnel) | 2-4 Week Half Life | Rich media home page unit on large portal – MSN or Yahoo |
| Online (lower funnel) | 1-2 Week Half Life | Media focused on in-market shoppers including retargeting |

Reference: https://www.fastcompany.com/1665084/how-long-does-your-ad-have-impact

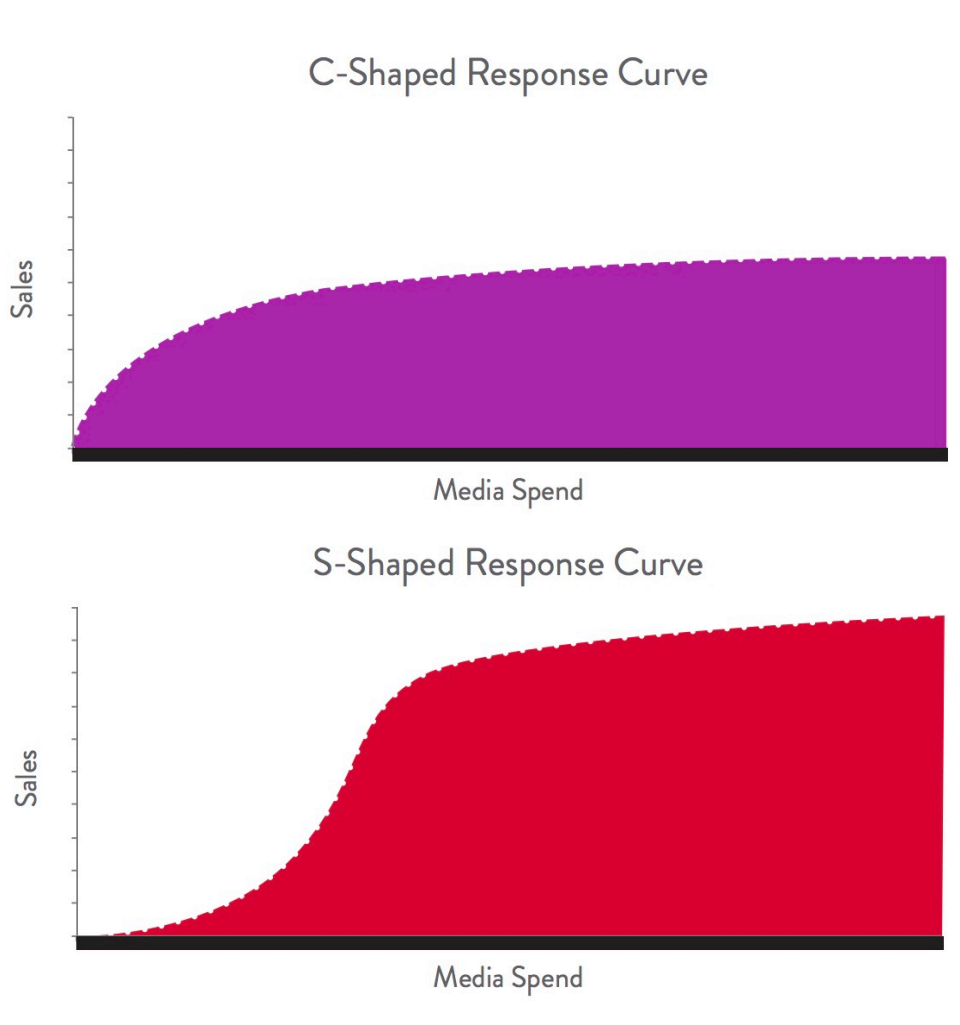
Copyright @ Data Application Lab

# *Dealing with Diminishing Returns*

- Several functions satisfy this property(Joy 2006), e.g, the log function and the logistic function.

$$y_t = \beta_0 + \beta_1 s_t + \varepsilon_t$$

$$s_t = \frac{1}{1 + e^{-\alpha x_t}} + \lambda s_{t-1}$$




Copyright @ Data Application Lab

# *Reason linear model might not perform well*

- Additive models imply **a constant absolute effect of each additional unit of explanatory variables**. They are suitable only if businesses occur in more stable environments and are not affected by interaction among explanatory variables. But in scenarios such as when pricing is zero, the sales (DV) will become infinite.
- To overcome the limitations inherent in linear models, **multiplicative models are often preferred**. These models offer a more realistic representation of reality than additive linear models do. In these models, IDVs are multiplied together instead of added.
- Changing the level of ad exposure brings about a relative change in sales volumes (**Shape Effect**).

Reference: https://www.latentview.com/marketing-mix-modeling/

Copyright @ Data Application Lab

# Model Selection

## Approach 2: Multiplicative regression, Log-Linear

**Semi-logarithmic models**

In Log-Linear models, the exponents of independent variables are multiplied.

$$Salest = exp(Intercept) * exp(\beta1*Pricingt) * exp(\beta2*Distributiont) * exp(\beta3*Mediat) * exp(\beta4*Discountst) * exp(\beta5*Seasonalityt) * exp(\beta6*Promotionst) *...$$

This can also be rewritten as

$$Salest = exp\ (Intercept + \beta1*Pricingt+ \beta2*Distributiont+ \beta3*Mediat+ \beta4*Discountst+ \beta5*Seasonalityt+ \beta6*Promotionst+ ...)$$

Logarithmic transformation of the target variable linearizes the model form, which in turn can be estimated as an additive model. The dependent variable is logarithmic transformed; the only difference between additive model and semi-logarithmic model.

$$Ln\ (Salest) = Intercept + \beta1*Pricingt+ \beta2*Distributiont+ \beta3*Mediat+ \beta4*Discountst+ \beta5*Seasonalityt+ \beta6*Promotionst+ ...$$

Some of the benefits of Log-Linear models are:

- The coefficients β can be interpreted as % change in business outcome (sales) to unit change in the independent variables.
- Each independent variable in the model works on top of what has been already achieved by other drivers. Hence, they are closer to real-time scenarios. A

Copyright @ Data Application Lab

Reference: https://www.latentview.com/marketing-mix-modeling/

# Model Selection

## Approach 3: Multiplicative regression, Log-Log

In Log-Log models, independent variables are also subjected to logarithmic transformation in addition to the target variable.

$$Salest = \exp(Intercept) * \beta1*Pricingt * \beta2*Distributiont * \exp(\beta3*Mediat) * \exp(\beta4*Discountst) * \exp(\beta5*Seasonalityt) * \exp(\beta6*Promotionst) *...$$

Rewriting the model in linear form,

$$Ln\ (Salest) = Intercept + \beta1*Ln\ (Pricingt)+ \beta2*Ln\ (Distributiont)+ \beta3*Mediat+ \beta4*Discountst+ \beta5*Seasonalityt+ \beta6*Promotionst+ ...$$

The main difference between Log-Linear and Log-Log models lies in the interpretation of response coefficients. In Log-Log models, the coefficients are interpreted as % change in business outcome (sales) in response to 1% change in independent variable

$$\beta = \%\Delta Dependent_Variable / \%\Delta Explanatory_Variable$$

This implies constant elasticity of the target variable to explanatory variables. In Log-Linear models, elasticity cannot be directly estimated but can be calculated from the coefficient as $\beta \cdot X$ for every time period. It increases in absolute value with the explanatory variable.

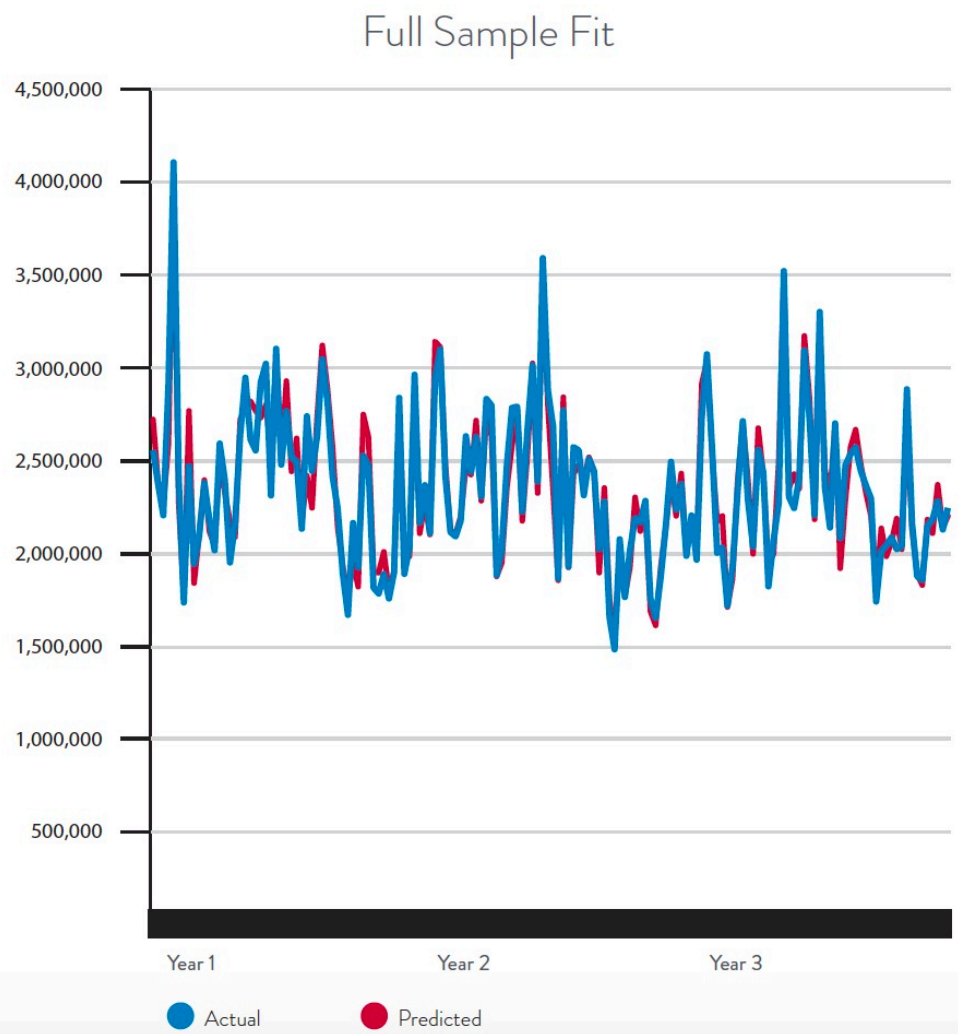
Copyright @ Data Application Lab

Reference: https://www.latentview.com/marketing-mix-modeling/

# *Step 4: Model Tuning*

- **R^2**
- **Mean Absolute Percentage Error (MAPE)**

| FULL SAMPLE FIT | TRAINING SAMPLE FIT | HOLDOUT SAMPLE FIT |
|---|---|---|
| 100% Sample | 80% Sample | 20% Sample |
| $R^2 = 0.93$<br>MAPE = 3.7% | MAPE = 3.8% | MAPE = 3.9% |



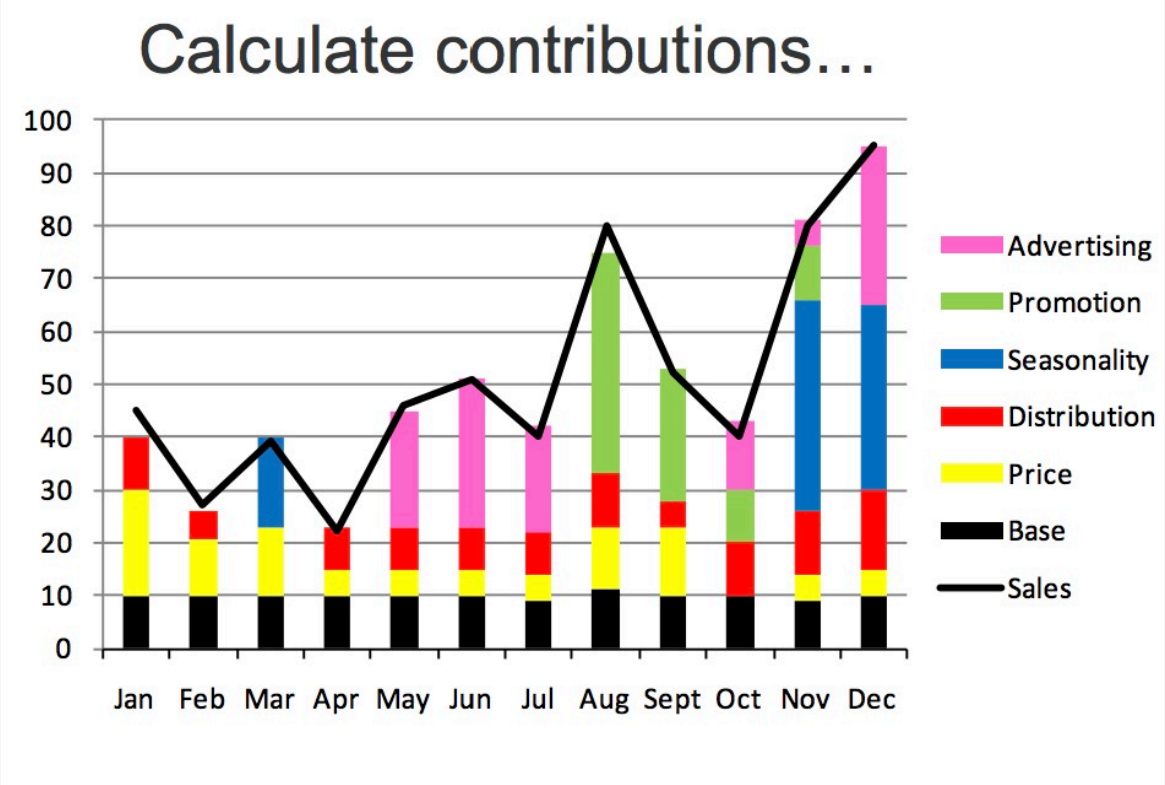

Copyright @ Data Application Lab

# *Step 5: Interpretation*

- Elasticity

| TRANSFORMATION | FUNCTION | ELASTICITY |
|---|---|---|
| Level-Level | $Y = a + bX$ | $\epsilon = b * \frac{X}{Y}$ |
| Log-Level | $log(Y) = a + bX$ | $\epsilon = b * X$ |
| Level-Log | $Y = a + b * log(X)$ | $\epsilon = \frac{b}{Y}$ |
| Log-Log | $log(Y) = a + b * log(X)$ | $\epsilon = b$ |

- Contribution




Copyright @ Data Application Lab

# Contribution

- Challenge
  - Distinguish the resulting sales into separate non-overlapping effect of different factors
- Solution
  - Contribution = Total Volume – Volume upon removal of the factor
- Example
  - $Y = \beta_0 * A^{\beta_1} * B^{\beta_2} = 1560$
  - $\beta_0 = 1000$
  - A = \$500, $\beta 1 = 0.03$, $A^{\beta_1} = 1.20$
  - B = \$200, $\beta 2 = 0.05$, $B^{\beta_2} = 1.30$ (B is more elastic than A)
  - $Y = \beta_0 A^{\beta_1} * B^{\beta_2} = 1000 * 1.20 * 1.30 = 1560$
  - Contribution of A = 1560 - 1560/1.20 = 260
  - Contribution of B = 1560 - 1560/1.30 = 360



Copyright @ Data Application Lab

# *Scaled Contribution*

- Challenge
  - Need to decompose Y(1560) to each factor
  - The contribution of each factor do not add up to total
  - $C_A + C_B = 260 + 360 = 620 > 560$
- Solution
  - Scale down each factor's contribution by removing additional volume proportionally
  - $C_i' = C_i - (Vp - Va) * C_i / Vp$, where $Vp = \sum C_i$
- Example
  - Scaled Contribution of A = 260 - (620 - 560) * 260/620 = 234.84
  - Scaled Contribution of B = 360 - (620 - 560) * 360/620 = 325.16
  - Total Volume = 1000 + 234.84 + 325.16 = 1560

$\beta_0 = 1000$
A = \$500, $\beta 1 = 0.03$, $A^{\beta 1} = 1.20$
B = \$200, $\beta 2 = 0.05$, $B^{\beta 2} = 1.30$ (B is more elastic than A)
$Y = \beta_0 A^{\beta 1} * B^{\beta 2} = 1000 * 1.20 * 1.30 = 1560$
Contribution of A = 1560 - 1560/1.20 = 260
Contribution of B = 1560 - 1560/1.30 = 360

Copyright @ Data Application Lab

# *ROI*

- ROI = Contribution / Marketing Investment
- ROI of A = Scaled Contribution of A / Spend of A

  = 234.84/500

  = 0.47
- ROI of B = Scaled Contribution of B / Spend of B

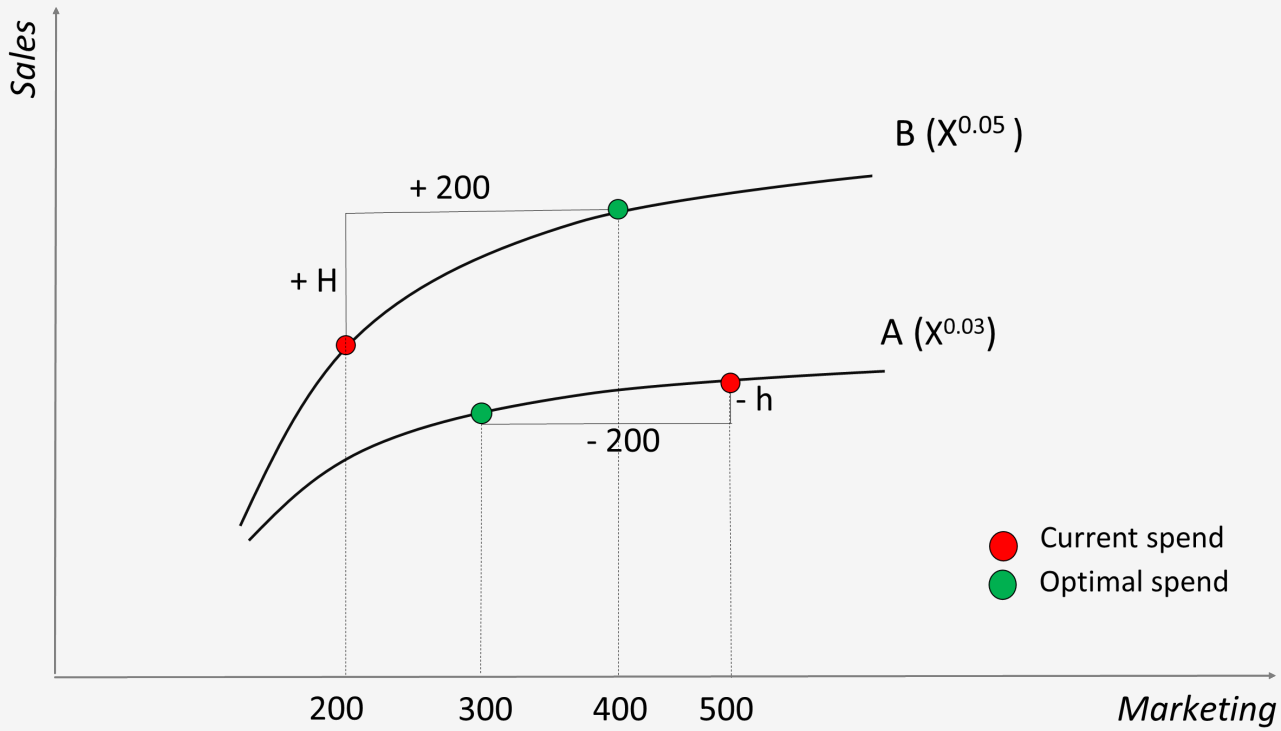
  = 325.16/200

  = 1.63
- B is more effective than A!



Copyright @ Data Application Lab

# Optimization




Move spending from A to B ;
Maximize H-h



Copyright @ Data Application Lab

# Optimization

- Before Optimization
  - A = $500
  - B = $200
  - $Y = \beta_0 A^{\beta_1} * B^{\beta_2} = 1000 * 500^{0.03} * 200^{0.05} = 1560$
- If move $200 from A to B
  - A = 500 - 200 = $300
  - B = 200 + 200 = $400
  - $Y = \beta_0 A^{\beta_1} * B^{\beta_2} = 1000 * 300^{0.03} * 400^{0.05} = 1601$
  - Gain form marketing optimization = 1601 -1560 =

$41 ≈ 2.63%



Copyright @ Data Application Lab

# Simulation & Optimization

| | | | |
|---|---|---|---|
| 450 | 250 | 700 | 1583.04542 |
| 440 | 260 | 700 | 1585.08387 |
| 430 | 270 | 700 | 1586.98288 |
| 420 | 280 | 700 | 1588.74933 |
| 410 | 290 | 700 | 1590.38919 |
| 400 | 300 | 700 | 1591.90762 |
| 390 | 310 | 700 | 1593.30905 |
| 380 | 320 | 700 | 1594.59724 |
| 370 | 330 | 700 | 1595.77534 |
| 360 | 340 | 700 | 1596.84595 |
| 350 | 350 | 700 | 1597.81114 |
| 340 | 360 | 700 | 1598.67246 |
| 330 | 370 | 700 | 1599.43098 |
| 320 | 380 | 700 | 1600.0873 |
| 310 | 390 | 700 | 1600.64153 |
| 300 | 400 | 700 | 1601.09329 |
| 290 | 410 | 700 | 1601.4417 |
| 280 | 420 | 700 | 1601.68536 |
| 270 | 430 | 700 | 1601.8223 |
| 260 | 440 | 700 | 1601.84996 |
| 250 | 450 | 700 | 1601.7651 |
| 240 | 460 | 700 | 1601.56374 |
| 230 | 470 | 700 | 1601.2411 |
| 220 | 480 | 700 | 1600.7914 |
| 210 | 490 | 700 | 1600.2078 |
| 200 | 500 | 700 | 1599.48216 |
| 190 | 510 | 700 | 1598.60482 |
| 180 | 520 | 700 | 1597.56428 |

Copyright @ Data Application Lab

# *ATTRIBUTION MODELS*

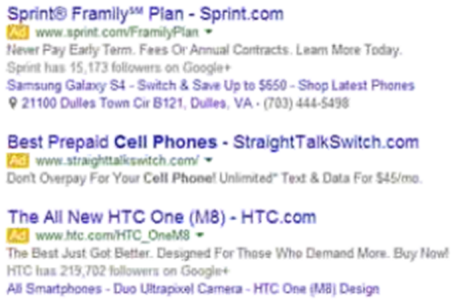
Copyright @ Data Application Lab

# Examples of Online Advertising

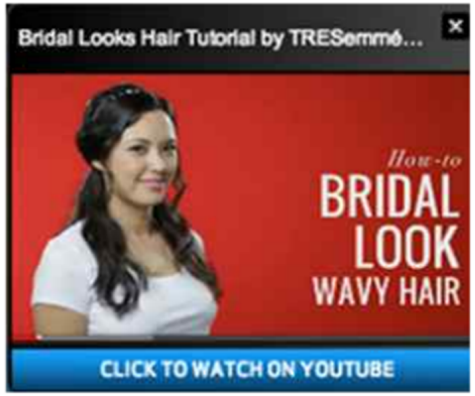
## Display Ad

## Email Ad

## Search Ads



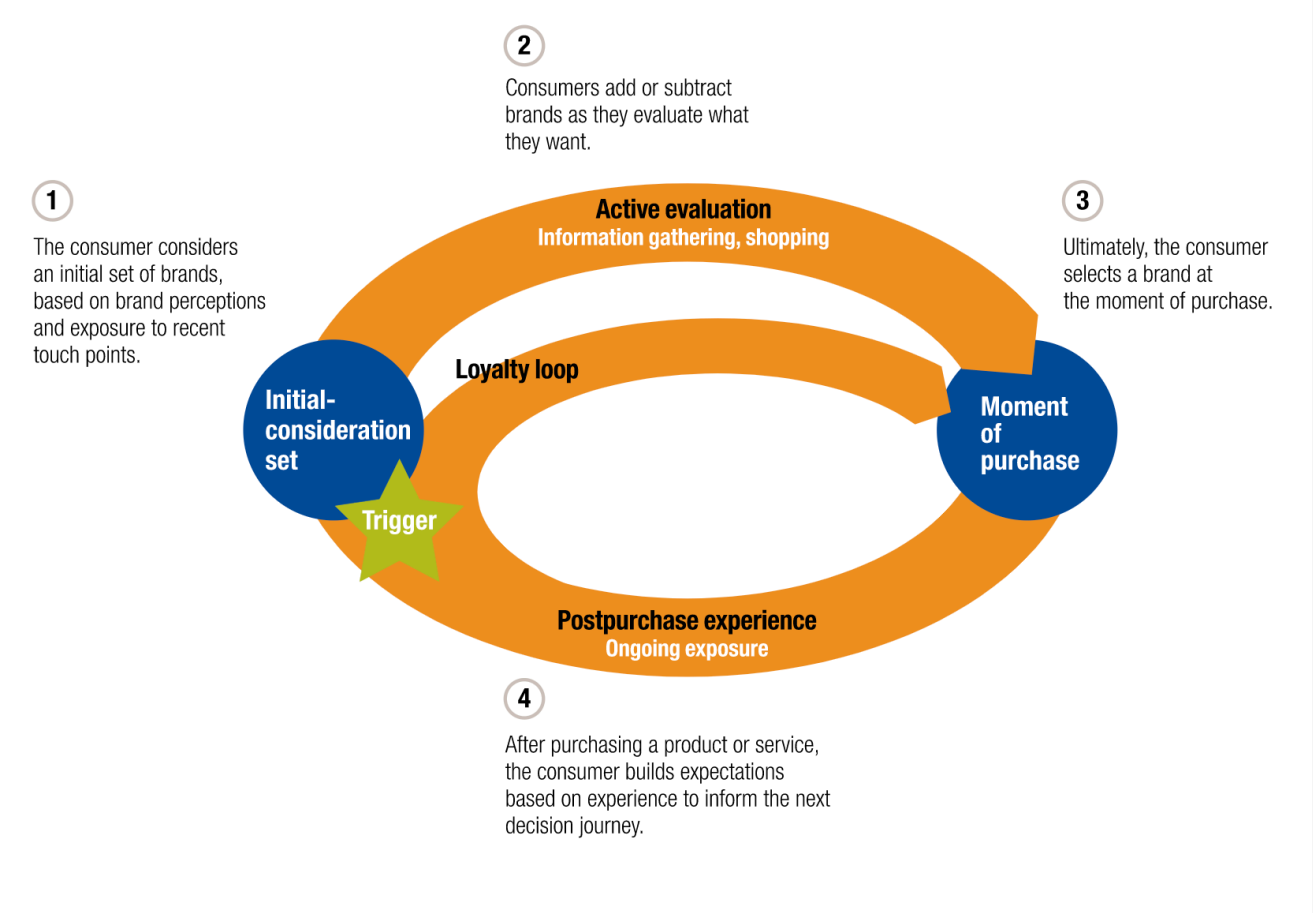

## Video Ads




## Social Media Ads

Copyright @ Data Application Lab

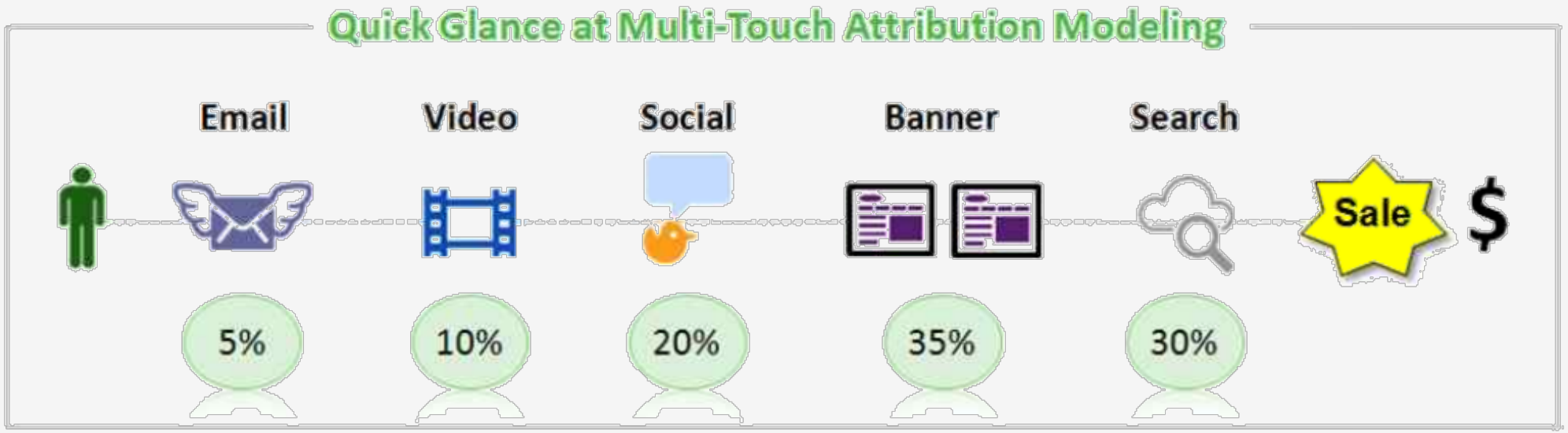
# Consumer's Side: Consumer Decision Journey




Copyright @ Data Application Lab

# Attribution

- Which TP gets credits for conversion?






Copyright @ Data Application Lab

# *Attribution Model*

Last Touch

Last Non-Direct Click

First Interaction

Linear Attribution

Time Decay

Position Based



Copyright @ Data Application Lab

# Last Touch

Customer Journey

| User | TP1 | TP2 | TP3 | TP4 | Conversion |
|---|---|---|---|---|---|
| 001 | Display | TV | Email | | 90 |
| 002 | Affiliate | Display | | | 100 |
| 003 | Video | Paid Search | Display | Social Media | 160 |
| 004 | TV | Paid Search | Video | | 180 |

Attribution

| TP | 001 | 002 | 003 | 004 | Total |
|---|---|---|---|---|---|
| Email | 90 | | | | 90 |
| Display | | 100 | | | 100 |
| Social Media | | | 160 | | 160 |
| Video | | | | 180 | 180 |



Copyright @ Data Application Lab

# Linear Attribution

Customer Journey

| User | TP1 | TP2 | TP3 | TP4 | Conversion |
|---|---|---|---|---|---|
| 001 | Display-30 | TV-30 | Email-30 | | 90 |
| 002 | Affiliate-50 | Display-50 | | | 100 |
| 003 | Video-40 | Paid Search-40 | Display-40 | Social Media-40 | 160 |
| 004 | TV-60 | Paid Search-60 | Video-60 | | 180 |

Attribution

| TP | 001 | 002 | 003 | 004 | Total |
|---|---|---|---|---|---|
| Display | 30 | 50 | 40 | | 120 |
| TV | 30 | | | 60 | 90 |
| Affiliate | | 50 | | | 50 |
| Video | | | 40 | 60 | 100 |
| Social Media | | | 40 | | 40 |
| Email | 30 | | | | 30 |
| Paid Search | | | 40 | 60 | 100 |



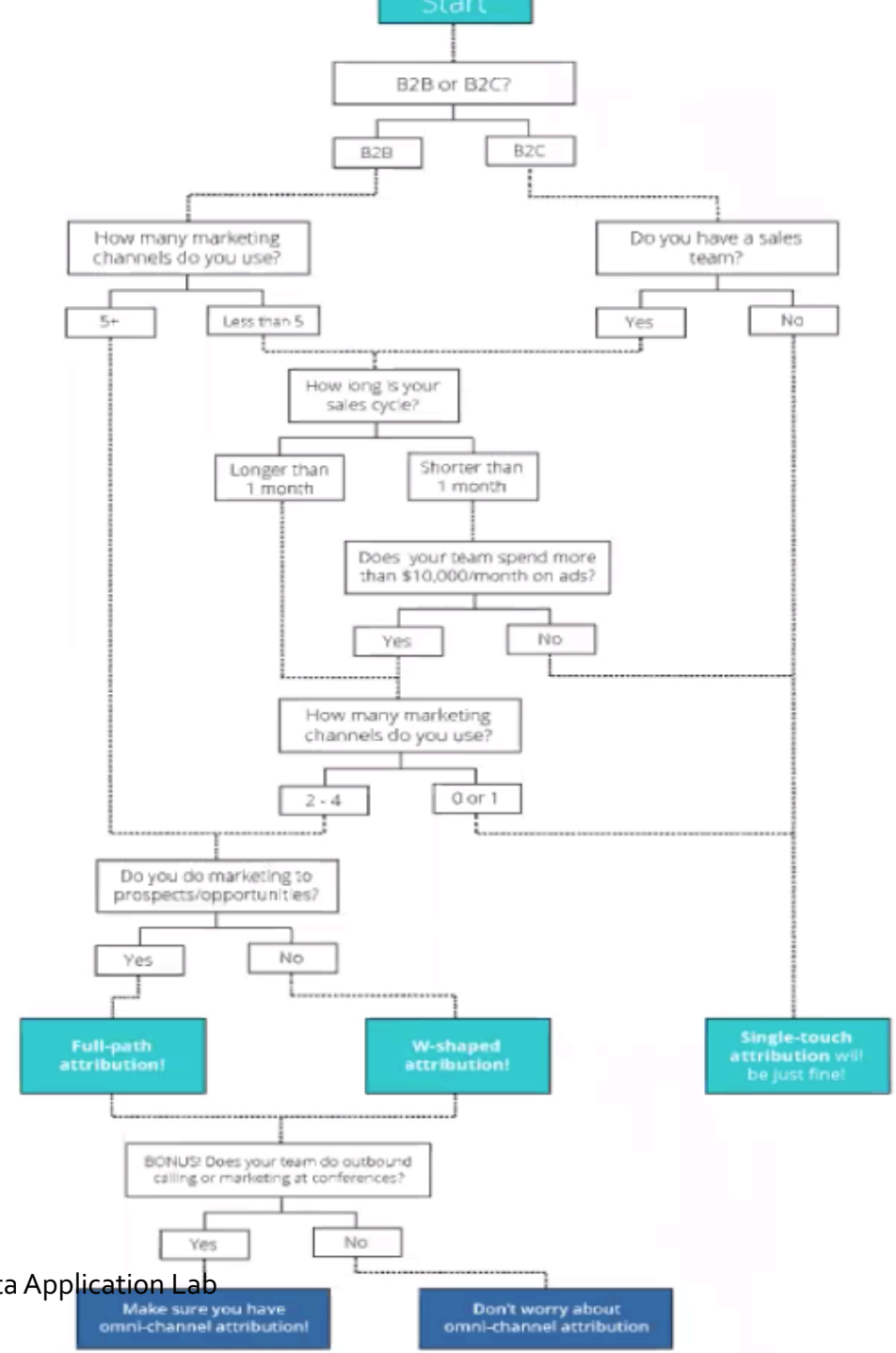
Copyright @ Data Application Lab

# *Choose attribution model*

- https://www.snapapp.com/blog/marketing-attribution-models/



Copyright @ Data Application Lab




# *Which attribute model is good?*



Copyright @ Data Application Lab

# *Digital Attribution vs MMM*

- Both are used to measure marketing effectiveness
- Often used together to help with Marketing Strategy and Budget Optimization
- Difference
  - Weekly aggregated vs. customer level
  - MMM: Top Down, all channels, all factors
  - Digital Attribution: Bottom Up, online channels



Copyright @ Data Application Lab

# CAUSAL INFERENCE

Copyright @ Data Application Lab

# *Approaches*

- A/B test (holdout)
- Forensic control

- More causal, more control
- Less elasticity information
- Sometimes harder to get enough sample size for some variables



Copyright @ Data Application Lab