Session 3: Marketing Research

Marketing Management
Nanda Kumar, Ph.D.

Agenda

- Marketing Research
  - Qualitative Analysis
  - Quantitative Analysis
- Quantitative Analysis
  - Regression Analysis
  - Analyzing survey data – factor analysis
  - Forecasting Product Diffusion – Bass Model

Product (Re)Positioning?

- How can we make our product better?
- Data Sources
  - Secondary – existing data
    - Internal – marketing department, sales, R&D etc.
    - External – trade organizations, business/trade/academic publications, company statements, govt. publications (census)
  - Primary – primarily collected for a specific purpose
Marketing Research

- Qualitative Analysis
  - Fair amount money spent
  - Often not comprehensive enough in itself
  - Can be:
    - Phenomenological
      - Understanding phenomena – how consumers buy, what they look for etc.
    - Exploratory
      - Gather information; may form the basis of quantitative analyses
    - Clinical
      - Rationale behind phenomena – why consumers behave the way they do?

Qualitative Research Methods

- Focus groups
- Observations
- Surveys
- Panels
- Experiments

Quantitative Research Methods

- Correlations - relationship between variables
- Review of Regression Analysis
- Survey analysis – Factor Analysis
- Forecasting Market Potential and Sales (Supplement)
Data: An Example

<table>
<thead>
<tr>
<th>Catalog</th>
<th>New</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>50</td>
</tr>
</tbody>
</table>

Relationship

Data

<table>
<thead>
<tr>
<th>Catalog</th>
<th>Old</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>B</td>
<td>500</td>
<td>500</td>
</tr>
</tbody>
</table>
### Relationship

![Graph showing relationship between new and old data percentages for buying A.](image)

### Data

<table>
<thead>
<tr>
<th>Catalog</th>
<th>New</th>
<th>Old</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>110,300</td>
<td>11,500</td>
</tr>
<tr>
<td>B</td>
<td>20,700</td>
<td>76,600</td>
</tr>
</tbody>
</table>

### Relationship

![Graph showing a negative trend between new and old data percentages for buying A.](image)
Correlation Coefficient ($r$)

- Statistical measure of the strength of relationship between two variables
- $r \in [-1,1]$ indicates a positive relationship
- $r \in [0,1]$ indicates a positive relationship
- $r \in [-1,0]$ indicates a negative relationship

Know your Data

- Sample should be representative of the population data
- Reason why experts advocate the use of random samples

Regression Analysis

- What does it do?
  - Uncovers the relationship between a set of variables
- Simple Regression
  - $y = f(x)$
- Regression sets out to find the $f(x)$ that best fits the data
Assumptions:

- \( f(x) \) is known up to some parameters
  - So \( f(x) = a + bx \)
  - Problem: Find \( a, b \) that best fit the data
- An Example:
  - \( Sales = a + b \times Price \)

How does it Work?

- Finds \( a, b \) that **best fit the data**
- Further assumptions:
  - \( Sales = a + b \times Price + \text{error} \)
  - Error is distributed normally: \( N(0, \sigma^2) \)
  - Criteria – finds \( a, b \) that minimize the sum of squared errors.
Hypothesis:
- Customers who purchase more frequently also buy bigger ticket items

Data

<table>
<thead>
<tr>
<th>Number of Purchases (X)</th>
<th>Largest Dollar Item (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>35</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
</tr>
</tbody>
</table>

Regression Model

- \( Y = a + b \times X + \text{error} \)
- Estimates: \( a = -18.22 \) \( b = 10 \)
- Goodness of Fit Measure: \( R^2 = 0.946 \)
Multiple Regression

- $Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_n X_n$
- Same as Simple Regression in principle
- New Issues:
  - Each $X_i$ must represent something unique
  - Variable selection

Example 1:
- $\text{Spending} = a + b \text{ income} + c \text{ age}$

Example 2:
- $\text{Sales} = a + b \text{ price} + c \text{ advertising} + d \text{ comp_price}$

Survey Analysis: Measurement of Department Store Image

- Description of the Research Study:
  - To compare the images of 5 department stores in Chicago area -- Marshal Fields, Lord & Taylor, J.C. Penny, T.J. Maxx and Filene’s Basement
  - Focus Group studies revealed several words used by respondents to describe a department store e.g. spacious/cluttered, convenient, decor, etc.
  - Survey questionnaire used to rate the department stores using 7 point scale
Items Used to Measure Department Store Image

1. Convenient place to shop
2. Fast checkout
3. Store is clean
4. Email is not well organized
5. Store is messy, cluttered
6. Convenient store hours
7. Store is far from home, school or work
8. Store has bad atmosphere
9. Attractive decor inside
10. Store is spacious

Pair-wise Correlations among the Items Used to Measure Department Store Image

<table>
<thead>
<tr>
<th></th>
<th>X_1</th>
<th>X_2</th>
<th>X_3</th>
<th>X_4</th>
<th>X_5</th>
<th>X_6</th>
<th>X_7</th>
<th>X_8</th>
<th>X_9</th>
<th>X_{10}</th>
</tr>
</thead>
<tbody>
<tr>
<td>X_1</td>
<td>1.00</td>
<td>0.79</td>
<td>0.41</td>
<td>0.26</td>
<td>0.12</td>
<td>0.09</td>
<td>0.87</td>
<td>0.37</td>
<td>0.32</td>
<td>0.18</td>
</tr>
<tr>
<td>X_2</td>
<td>1.00</td>
<td>0.32</td>
<td>0.21</td>
<td>0.20</td>
<td>0.90</td>
<td>0.85</td>
<td>0.31</td>
<td>0.35</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>X_3</td>
<td>1.00</td>
<td>0.00</td>
<td>0.76</td>
<td>0.34</td>
<td>0.40</td>
<td>0.82</td>
<td>0.78</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X_4</td>
<td>1.00</td>
<td>0.75</td>
<td>0.30</td>
<td>0.28</td>
<td>0.78</td>
<td>0.26</td>
<td>0.39</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X_5</td>
<td>1.00</td>
<td>0.76</td>
<td>0.30</td>
<td>0.39</td>
<td>0.39</td>
<td>0.76</td>
<td>0.75</td>
<td>0.76</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>X_6</td>
<td>1.00</td>
<td>0.29</td>
<td>0.26</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X_7</td>
<td>1.00</td>
<td>0.82</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X_8</td>
<td>1.00</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X_{10}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>
Factor Analysis for the Department Store Image Data:
Variance Explained by Each Factor

<table>
<thead>
<tr>
<th>Factor (Latent Root)</th>
<th>Variance Explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>5.725</td>
</tr>
<tr>
<td>Factor 2</td>
<td>2.761</td>
</tr>
<tr>
<td>Factor 3</td>
<td>0.366</td>
</tr>
<tr>
<td>Factor 4</td>
<td>0.357</td>
</tr>
<tr>
<td>Factor 5</td>
<td>0.243</td>
</tr>
<tr>
<td>Factor 6</td>
<td>0.212</td>
</tr>
<tr>
<td>Factor 7</td>
<td>0.132</td>
</tr>
<tr>
<td>Factor 8</td>
<td>0.123</td>
</tr>
<tr>
<td>Factor 9</td>
<td>0.079</td>
</tr>
<tr>
<td>Factor 10</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Perceptual Map

F1 - Convenience

F2 - Ambience

Product Positioning & Perceptual Maps

- Information Needed for Positioning Strategy:
  - Understanding of the dimensions along which target customers perceive brands in a category and how these customers perceive our offering relative to competition
  - How do our customers (current or potential) view our brand?
  - Which brands do those customers perceive to be our closest competitors?
  - What product and company attributes seem to be most responsible for these perceived differences?
- Competitive Market Structure
  - Assessment of how well or poorly our offerings are positioned in the market
Product Positioning & Perceptual Maps (cont.)

- **Managerial Decisions & Action:**
  - Critical elements of a differential strategy/action plan
  - What should we do to get our target customer segment(s) to perceive our offering as different?
  - Based on customer perceptions, which target segment(s) are most attractive?
  - How should we position our new product with respect to our existing products?
  - What product name is most closely associated with attributes our target segment perceives to be desirable?

Estimating Market Potential

- Estimate number of potential buyers
  - Purchase intention surveys
  - Extrapolate to the set of potential buyers
  - Deflate the estimates ~ factor of 2
- Estimate purchase rate
  - Market potential = Potential buyers * purchase rate

Forecasting Sales: Bass Model

- Basic Idea:
  - Probability that a customer will purchase at time t conditional on not having purchased until that time = p + q*(number of customers bought so far)
  - Hazard rate = p + q*Cumulative Sales
  - Solution yields an expression for Sales(t) = g(p,q,m)
Forecasting Sales

- Need a few observations
- Step 1: Estimate market potential $m$
- Step 2: Empirically estimate $p$ and $q$
- Step 3: Given $p$, $q$ and $m$ – Sales($t$) from Bass model, simulate Sales($t$) by varying $t$

Example

Adoption of Answering Machines, 1982-1993

Year Fitted Adoption

Another Example 35 mm Projectors

Actual and Fitted Adoption of 35 mm Projectors, 1965-1986,

Year Fitted
Another Example: Overhead Projectors

Marketing Research: Supplement

Regression Analysis

Diagnostics

- Linearity Assumption
  - Y is linear in X – does this hold?
  - If not transform the variables to ensure that the linearity assumption holds
  - Common Transforms: Log, Square-root, Square etc.
Plot Y vs. X ($r=0.97$)

![Y vs X plot]

Plot $Y^{1/2}$ vs. X ($r=0.99$)

![Sqrt(Y) vs X plot]

Regression Model

- $Y^{1/2} = a + b\ X + \text{error}$
- Estimates: $a = 0.108845\ b = 0.984$
- Goodness of Fit Measure: $R^2 = 0.9975$
Obsession with $R^2$

- Can be a misleading statistic
- $R^2$ can be increased by increasing the number of explanatory variables
- $R^2$ of a bad model can be higher than that of a good model (one with better predictive validity)

Marketing Research: Supplement

Bass Model

Forecasting Sales: The Bass Model

- $f(t)/[1-F(t)]=p+qF(t)$  Hazard Model
- $m=$ ultimate market potential
- $p=$ coefficient of innovation
- $q=$ coefficient of imitation
- $S(t)=mf(t)=m[p+qF(t)][1-F(t)]$  
  $=pm+(q-p)Y(t)-(q/m)y(T)^2$
A Differential Equation

- Solution: $S(t) = \frac{m[(p+q)^2/p]e^{(p+q)t}}{1+(q/p)e^{(p+q)t}}^2$
- $t^* = \frac{1}{p+q} \ln(q/p)$

$t^*$ = Time of Peak Sales

- Beautiful!