Shidler College of Business
University of Hawai`i
Business Plan Competition
Bootcamp Workshop
2010

Professor Dana L. Alden
February 6, 2010
Why Marketing Research Is So Crucial to Long-term Success?

- Avoids costly mistakes:
- Pillsbury Cake Mix
- New Coke
- Provides a check on theory, experience and intuition
- Provides Continuous Feedback from Customers
- Promotes Competitive Advantage in World of Rapid Response
Marketing Research Process

Initial Problem Definition →

Exploratory Research →

Problem Refinement →

Study Design →

Data Collection →

Data Analysis →

Data Dissemination

Dana L. Alden, Ph.D.
Basic Types of Customer Research

- Exploratory, Descriptive, Experimental
- Primary versus Secondary
- Observation versus Questioning
Secondary Data

- Syndicated data are data that are collected in a standard format and made available to all subscribers. Both B2B and B2C.

Example: The Nielsen TV Ratings
http://www.nielsenmedia.com/nc/portal/site/Public/
Dun & Bradstreet credit bureau collects vast amount of information on business firms, private and public.

- 4 million firms in data base updated monthly
- Whereas SIC uses 4 digits and NAICS uses 5 or 6, DMI’s uses 8 digits to classify firms into very specific types of businesses.
Dun’s Market Identifiers (continued)

- Manufacturer of wooden gift baskets
- Small without marketing resources
- Needed wholesalers…of gift baskets. Where are they? Who are they?
- SIC: 51 Code for wholesalers/nondurables
- 5199, wholesalers, nondurables, miscellaneous goods! Thousands of firms wholesaling all kinds of goods.
- 51990603 DMI code: wholesalers of gift baskets. 45 in U.S. Complete info on each!

Dana L. Alden, Ph.D.
Secondary Data Examples

B2C

- Measuring Consumer Attitudes and Opinion Polls
  - Yankelovich Monitor
  - Harris Poll
  - Gallup Poll

- Defining Market Segments
  - Dun’s Market Identifiers (DMI) for b2b
  - VALS
Secondary Data Example
Gallup.com

- Public opinion polling
- Wide variety of topics: Iraq, military & defense, stem cell research, important factors in buying a car, smoking population percents over time, guns, etc.
- Some issues may be tracked over several years; Gallup started in 1935.
- [http://www.gallup.com](http://www.gallup.com) go to “Consumers”

Dana L. Alden, Ph.D.
Claritas has a market profiling system called PRIZM.
- Defines every neighborhood in the U.S. in terms of 66 different clusters.
- ESRI’s Community Tapestry
- 65 segments

http://www.claritas.com/MyBestSegments/Default.jsp
Secondary Data Examples
Scanner Data

– Market tracking studies monitor, or track, a variable over time.
  - ACNielsen also has Convenience Track® for tracking products sold through c-stores
  - Infoscan’s Custom Store Tracking 34,000 supermarkets, drugstores, and mass merch. – weekly
  - IRI Consumer Network Panel
HOUSEHOLD DATA are collected through panels that scan at home.
Secondary Data Examples
Radio Listenership, Magazine Readership, Media Habits

- Monitoring Media Usage and Promotion Effectiveness
  - Radio – Arbitron
    - Shifting to Portable People Meter

- Print – NOP World’s Starch Readership Service

- Multimedia – Simmons National Consumer Study
Using Secondary Data

Estimating Demand – Chain Ratio Method: Light Beer in Hawaii

- (500,000 people > 21 x $50k per capita avg. inc.) x
- (25% discretionary income spent on food) x
- (10% of food $ spent on beverages) x
- (10% of beverages $ spent on beer) x
- (40% of beer $ spent on light beer) =
- $25 million market potential

Dana L. Alden, Ph.D.
**Estimating Demand - Market Build Up Method:**

**Mid-sized 4-Door Car Sales in Bangalore, India**

- Given knowledge of Target Market, estimate number of families in Bangalore with incomes $40K -$75K
- Use National Data to determine numbers of mid-size, 4-Door Cars sold to such families for all India
- % families in Bangalore x Total midsize, 4-Door Cars Sold in India =
- Area Potential Car Sales
“Survey of Buying Power”

- The SBP is an annual survey published in Sales and Marketing Management magazine.
- The “Survey of Buying Power” provides the buying power index (BPI).

Dana L. Alden, Ph.D.
“Calculating the Buying Power Index”

\[
BPI = 0.20(\text{Population of Market Area A/ Total U.S. Population}) + 0.50(\text{EBI of Market Area A/Total U.S. EBI}) + 0.30(\text{Retail Sales of Market Area A/Total U.S. Retail Sales})
\]

BPI (your town) = 0.20(1%) + 0.50(1.2%) + 0.30(1.3%) = 1.19%

So … 1.19% of media budget should go to your town.

- Scroll down to “5 Year Projections” then, Buying Power Index.
### Example of Buying Power Index: Three Urban Areas

<table>
<thead>
<tr>
<th>County</th>
<th>% Income</th>
<th>% Retail sales</th>
<th>% Households</th>
<th>% Exp. Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cook</td>
<td>.5(2.2%) + .3(2.5%) + .2(1.6%) = 2.17%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. A.</td>
<td>.5(2.5%) + .3(2.5%) + .2(2.3%) = 2.46%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.Y.</td>
<td>.5(2.3%) + .3(2.8%) + .2(2.0%) = 2.39%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dana L. Alden, Ph.D.
**Primary Data: Exploratory**

**TIPS for Successful Focus Groups**

- Know f.g. limitations and clear research objectives
- Recruit appropriately
- Select professional moderator carefully
- Work-out discussion guide and follow
- No thought is “dumb”
- Conduct in professional setting

Dana L. Alden, Ph.D.
Primary Data: Descriptive
Minimizing Survey Response Problems
Through Effective Question Design

1. Be Specific - avoid questions that are too broad.

   Q. In the past six months, has your household purchased any new appliances from the store?

   Q. I have a list of household appliances. As I read each one, please tell me whether or not your household has purchased this type of appliance new from the store during the past six months? How about …
   a. A refrigerator  ___Yes ___ No
   b. A kitchen range or oven  ___ Yes ___ No

2. Specify who, what, when, where and how.

   Q. In 2009, about what was your total family income before taxes? Please count income from all members of your household and from all sources, including sources such as interest and dividends.

Dana L. Alden, Ph.D.
Minimizing Survey Response Problems Through Question Design

3. Avoid “leading” or “loaded” questions.
Q. Shouldn’t concerned parents use car seats?
Q. Research shows that children are safest when they are required to sit in the rear seat of cars until they are at least 14. How likely are you to require your child under 14 to sit in the rear seat?

Q. Please rate the quality of the food and the service at our restaurant.
Q. Please rate the quality of the food at our restaurant.

5. Use simple language and short questions.
Q. If the Sears repair service schedule was not convenient for you, would you consider or not consider calling a competing repair organization to fix the problem you have?
Q. How likely are to use repair services other than Sears for major appliance repair?
6. Use numbers rather than indefinite adjectives to measure magnitudes, e.g., instead of asking whether someone purchases a product “regularly” use specific numbers per week or month.

7. Avoid words with double and/or slang-oriented meanings, e.g., “bad,” “cool.”

8. Avoid words that are extreme, e.g., always, never, everyone, any, anybody, ever

9. Pretest to see whether respondents have trouble understanding some of the questions.

10. Pretest to learn what respondents have in mind when they answer key questions.

Dana L. Alden, Ph.D.
Question Flow in a Questionnaire

- Opening questions – *Warm ups*
- Prior to major sections – *Transitions*
  
  You’re half-way there! Now a few questions about how you rate the quality of our business.

- Middle of questionnaire - *Difficult-to-Answer*
  
  – Rate each of the following aspects of Foodland on how satisfactory it is for you.

- Last Section - *Classification*
  
  – What is the highest level of education you have earned?
### Examples of Data Collection Trade-Offs

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Focus Groups</th>
<th>Personal Interviews</th>
<th>Telephone Interviews</th>
<th>Mail/Internet Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>low</td>
<td>high</td>
<td>moderate</td>
<td>low</td>
</tr>
<tr>
<td>Flexibility</td>
<td>high</td>
<td>high</td>
<td>moderate</td>
<td>low/high</td>
</tr>
<tr>
<td>Int. Bias Potential</td>
<td>high</td>
<td>high</td>
<td>moderate</td>
<td>low</td>
</tr>
<tr>
<td>Sample Bias Potential</td>
<td>high</td>
<td>moderate</td>
<td>low</td>
<td>high</td>
</tr>
</tbody>
</table>

Dana L. Alden, Ph.D.
Understanding Data
Via Descriptive Analysis

- Two sets of descriptive measures:
  - Measures of central tendency: used to report a single piece of information that describes the most typical response to a question
  - Measures of variability: used to reveal the typical difference between the values in a set of values
Understanding Data Via Descriptive Analysis

• Measures of Central Tendency:
  – **Mode**: the value in a string of numbers that occurs most often
  – **Median**: the value whose occurrence lies in the middle of a set of ordered values
  – **Mean**: sometimes referred to as the “arithmetic mean”; the average value characterizing a set of numbers

\[
\text{Mean (}\overline{X}\text{)} = \frac{\sum_{i=1}^{n} x_i}{n}
\]

where:
- \( n \) = the number of cases
- \( x_i \) = each individual value
- \( \sum \) signifies that all the \( x_i \) values are summed

Dana L. Aiden, Ph.D.
Understanding Data Via Descriptive Analysis

- Measures of Variability:
  - **Frequency distribution** reveals the number (percent) of occurrences of each number or set of numbers
  - **Range** identifies the maximum and minimum values in a set of numbers
  - **Standard deviation** indicates the degree of variation in a way that can be translated into a bell-shaped curve distribution

\[
(s) = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}}
\]

Dana L. Alden, Ph.D.
## What Descriptive Statistic to Use When

<table>
<thead>
<tr>
<th>Example Question</th>
<th>Measurement Level</th>
<th>Central Tendency (the most typical response)</th>
<th>Variability (how similar the responses are)</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is your gender?</td>
<td>Nominal scale</td>
<td>Mode</td>
<td>Frequency and/or percentage distribution</td>
</tr>
<tr>
<td>Rank these 5 brands from your 1st choice to your 5th choice</td>
<td>Ordinal scale</td>
<td>Median</td>
<td>Cumulative percentage distribution</td>
</tr>
<tr>
<td>On a scale of 1 to 5, how does &quot;Starbucks&quot; rate on variety of its coffee drinks?</td>
<td>Interval scale</td>
<td>Mean</td>
<td>Standard deviation and/or range</td>
</tr>
<tr>
<td>About how many times did you charge your cell phone last week?</td>
<td>Ratio scale</td>
<td>Mean</td>
<td>Standard deviation and/or range</td>
</tr>
</tbody>
</table>
Parameter Estimation

Parameter estimation: the process of using sample information to compute an interval that describes the range of values of a parameter such as the population mean or population percentage is likely to take on
Parameter Estimation

- Parameter estimation involves three values:
  1. **Sample statistic** (e.g., the mean)
  2. **Standard error** (variance divided by sample size; formula for standard error of the mean = $s/\sqrt{n}$)
  3. **Confidence interval** (gives us a range within which the mean will fall if we were to repeat the study many times over)
Parameter Estimation

- Confidence intervals:
  - the degree of accuracy desired by the researcher and stipulated as a level of confidence

- Most commonly used level of confidence:
  - 95%; corresponding to 1.96 standard errors
Parameter Estimation

- What does this mean? It means that we can say that if we did our study over 100 times, we can determine a range within which the sample statistic will fall 95 times out of 100 (95% level of confidence). This gives us confidence that the real population value falls within this range.
How do I interpret the confidence interval?

- Take many, many, many samples
- Plot the means
- 95% will fall in confidence interval
Parameter Estimation: Mean

From a restaurant data set assume
- We want to know how much people living in Honolulu are likely to spend on upscale dining in a month – a metric variable in dollars
- Say we have a random sample of 500
- Say we want to estimate 99% confidence where \( z = 2.58 \)
- Say the sample mean is $65.00
- Say standard deviation is $20.00
- Then the population parameter, \( \mu \), is estimated using the sample mean
  - $65 \pm \left( (2.58) \times \left( \frac{s}{\sqrt{n}} \right) \right)$
  - $65 \pm \left( (2.58) \times \left( \frac{20}{22.36} \right) \right) = 65 \pm 2.31$
  - $62.69 \leq \mu \leq 67.31$
Estimating a Population Mean

- How do we interpret the results?
  - My best estimate is that people in Honolulu spend $65 per month dining out at upscale restaurants.
  - I am 99% confident that the true population mean falls between $62.69 and $67.31.
  - 99 samples out of 100, we will find a sample mean between $62.69 and $67.31.
Netflix wants to explore in-depth opportunities and problems consumers perceive with using its entertainment service.

Bank of Hawai’i wants to measure service quality perceptions of its customers over time.

P&G introduces a new detergent in Japan & wants to measure change in brand recognition (“share of mind”) now vs 1 year later.

Toyota wants to develop a media strategy for launch of new, mid-priced sports car to younger drivers.

McDonald’s wants to know whether a soy burger sandwich that has done well in lab tests will do well in the SF and LA before launching in those cities.
Hawaii-based Marketing Research Providers

- SMS Research:
  Hersh Singer/Jim Dannemiller
  http://www.smshawaii.com/

- Omnitrak Group, Inc.:
  Pat Loui - 528-4050

- QMark Research and Polling:
  Barbara Ankersmitt - 524-5194

- Ward Research:
  Rebecca Ward - 522-5123
Any Questions?