MARKET SEGMENTATION SUMMARY

• The purpose of marketing is to match the needs of consumers with the offers of suppliers that can satisfy those needs which benefits both parties.

• Marketing planning involves setting objectives and formulating plans to achieve them.

• A marketing plan has two components: a strategic plan outlining the long-term direction and a tactical plan translating the strategic plan into short-term action.

• The strategic plan states where the organization wants to go while the tactical plan contains instructions on how to get there.

• Before climbing a mountain, it is important to determine the present location and then decide which mountain to climb. This is the strategic decision that determines subsequent tactical decisions.

• Preparations for marketing are similar in that the strategic plan identifies consumer needs, organizational strengths and weaknesses, and external opportunities and threats.

• The text discusses strategic and tactical marketing planning. Strategic marketing planning involves deciding which consumers to focus on (segmentation and targeting), and which image to create in the market (positioning).

• A SWOT analysis identifies an organization's strengths, weaknesses, opportunities and threats to outline what the organization can offer consumers. Market research explores consumer needs and desires.

• Tactical marketing planning covers product development, price determination, distribution channels selection, and promotion. It depends on the strategic marketing plan but not vice versa.

• Strategic marketing identifies the suitable "mountain to climb" while tactical marketing determines the equipment, route, pace and people for the climb.

• Strategic marketing refers to the overall direction and vision of an organization's marketing efforts. It focuses on the big picture and long-term goals.

• Tactical marketing involves the specific activities and campaigns used to implement the strategic marketing plan. It focuses on short-term goals and objectives.

• Good strategic marketing is essential for organizational success. Even excellent tactical marketing cannot compensate for a bad strategic marketing plan.

• The combination of good strategic and tactical marketing leads to the best outcomes for an organization.

• Market segmentation is used by marketers to select target markets and design marketing mixes for products.

• Market segmentation divides a broad market into smaller segments based on consumer characteristics and needs.

• Successful firms drive their businesses based on segmentation. It is critical to marketing success.

• Smith (1956) proposed market segmentation as a strategy to view a heterogeneous market as smaller homogeneous markets.

• The key is to segment consumers into groups that are similar within the segment but different across segments.

• Segmentation criteria can be a single characteristic like age, gender, lifestyle or a set of characteristics like benefits sought, values, expenditure patterns.

• The ideal segmentation has distinct segments that differ in needs and willingness to pay for product features. This allows companies to better satisfy the needs of each segment.

• Market segmentation forces organizations to reflect on where they stand and want to be in the future. It helps gain insights into what consumers want.

• Market segmentation leads to benefits like better understanding of consumer differences which improves the match between organizational strengths and consumer needs.

• An improved match can form the basis of a long-term competitive advantage in the selected target segments.

• Extreme competitive advantage results from being able to cater to the needs of a very specific niche segment. Ideal niches match the organization's skills, are profitable, have growth potential, and are different from competitors.

• Market segmentation involves dividing the market into distinct groups of consumers who have similar needs and wants. This allows companies to target their products and services more efficiently.

• Taking market segmentation to the extreme means offering customized products and services to very small groups of consumers, also known as micro marketing or hyper-segmentation.

• A marketing mix developed for a specific market segment is likely to yield a higher return on investment since less effort is wasted on consumers whose needs cannot be satisfied.

• For small businesses, market segmentation can be essential for survival since they lack resources to serve a larger market.

• Market segmentation has been shown to improve sales management by allowing sales efforts to be targeted at specific consumer groups.

• Market segmentation can contribute to team building within an organization since conducting segmentation requires representatives from different units to work together.

• Implementing market segmentation requires substantial investments of time, human resources, and finances. An unsuccessful segmentation strategy can waste these resources and fail to provide a competitive advantage.

• Market segmentation analysis involves grouping consumers into segments based on similar product preferences or characteristics.

• It is an exploratory and statistical process where many decisions by the analyst affect the final segmentation solution.

• Involving both data analysts and users who understand the organization's mission helps ensure high quality segmentation.

• Additional tasks like collecting good data, exploring the data, and profiling segments after extraction help improve segmentation.

• The data collection and segment extraction processes cannot compensate for poor quality data. The segmentation can only be as good as the input data.

• The text discusses conducting effective market segmentation analysis. It involves extracting market segments, collecting data and exploring it, profiling and describing segments.

• There are three layers to market segmentation analysis - the core tasks of segment extraction, segment profiling and description; the organizational implementation issues; and monitoring effectiveness over time.

• The organizational implementation issues include deciding whether segmentation will create opportunities, committing to the long-term strategy, collecting relevant data, selecting target segments, and developing a marketing plan for those segments.

• While data analysts can provide facts about market segments, organizational users make key decisions like selecting target segments and designing marketing mixes.

• There are multiple approaches to market segmentation analysis based on the organization's constraints and the variables used.

• Based on organizational constraints, there are three approaches: segment revolution, segment evolution, and segment discovery.

• Segment revolution involves a quantitative survey-based approach and requires the most radical change in the organization. It starts from scratch.

• Segment evolution is a less radical approach that refines an existing segmentation.

• Segment discovery involves finding existing segments and targeting them without a deliberate segmentation analysis.

• The quantitative survey-based or segment revolution approach is considered the prototypical market segmentation analysis but is not always viable in practice due to organizational unwillingness or inability to change sufficiently.

• Market segmentation analysis needs to be refined and sharpened over time through segment evolution rather than radical changes. Organizations can workshop to guide this process.

• Segments can also be discovered through exploratory research and data mining of large data streams, pointing to the need to adapt segmentation strategies.

• Segmentation approaches can use single or multiple variables, like age or expenditure patterns.

• Using a single segmentation variable is called a priori, convenience-group, or commonsense segmentation. It relies on intuition, secondary data analysis, or existing segments.

• Multiple segmentation variables allow for a multidimensional statistical analysis to identify segments with similar characteristics across variables.

• Segmentation variables can include age, expenditure patterns, lifestyles, interests and opinions.

• The text discusses different approaches to market segmentation analysis namely a priori segmentation, commonsense segmentation, and convenience-group segmentation.

• A priori segmentation means the decision about what characterizes each segment is made in advance before data analysis.

• Commonsense segmentation implies that market segments are chosen based on the providers' understanding of appropriate segments to target. The aim is to gain insight into segments rather than identify key segment characteristics.

• An example of commonsense segmentation is brand segmentation. However, identifying powerful segmentation variables can make commonsense segmentation more efficient.

• The proactive approach that exploits multiple segmentation variables is referred to as a posteriori, cluster based, or data-driven segmentation. This means the nature of market segments is known only after data analysis.

• Commonsense and data-driven segmentation are the two main approaches to market segmentation analysis.

• Data-driven segmentation relies on primary research and data analysis to identify market segments while commonsense segmentation uses logical grouping variables.

• In reality, most segmentation studies use a combination of both approaches sequentially or simultaneously.

• Multi-stage segmentation uses multiple segmentation variables in stages to identify target segments.

• When conducting data-driven segmentation, analysts assume naturally existing market segments can be identified in the data.

• However, naturally distinct segments rarely exist in consumer data. This raises the question of whether segments should be imposed if they do not naturally exist.

• Pioneers of market segmentation worked under the assumption that taxonomic procedures identify natural groups in the data.

• The text discusses data structure and data-driven market segmentation approaches.

• It explains four combinations of segmentation approaches based on the nature of segmentation variables used: commonsense/commonsense segmentation, commonsense/data-driven segmentation, data-driven/commonsense segmentation, and data-driven/data-driven segmentation.

• The primary segmentation variables can be commonsense variables like age and country of origin, or data-driven variables like expenditures and vacation activities. The secondary segmentation variables can also be commonsense or data-driven.

• Initially, the aim of market segmentation was to identify natural groupings in the data.

• More recently, market segmentation is seen as the process of creating artificial segments that can help develop effective marketing strategies, since empirical data sets typically do not show clear cluster structure.

• While early works aimed to identify natural segments, they acknowledged that artificial segments can be valuable for developing marketing strategies.

• The text discusses different conceptual approaches to market segmentation: natural, reproducible, and constructive segmentation.

• Natural segmentation assumes distinct market segments exist in the data and the aim is to find them. This is the traditional view.

• Reproducible segmentation assumes natural segments do not exist but the data has some structure, allowing similar segments to be found repeatedly. This makes results more reliable.

• Constructive segmentation assumes no data structure exists to find similar segments repeatedly. However, targeting subgroups can still be useful even if consumer preferences are spread evenly.

• It is often not known at the start which type of segmentation the data permits. Analysis found natural segmentation is extremely rare, occurring in only 6% of datasets studied. This means data structure analysis is essential before extracting market segments.

• Conducting data structure analysis in advance of the actual data-driven market segmentation analysis is recommended. It gives an overall picture of the data and helps avoid bad decisions.

• Repeatedly segmenting the data with different numbers of segments and algorithms can give insight into the data structure.

• The solution ranges from natural clustering where the analyst needs little user input, to unstructured data where the analyst must work closely with users to construct useful segments.

• The recommended 10 step approach for market segmentation analysis includes:

- Weighing the pros and cons of segmentation

- Specifying ideal segment characteristics

- Collecting or compiling data

- Exploring the data

- Extracting market segments from the data

- Profiling the segments

- Describing the segments in detail

- Selecting target segments

- Developing a customized marketing mix

- Evaluating and monitoring segments for changes

• Decide whether the market is suitable for segmentation and if you can make a long-term commitment to it.

• Specify your ideal target segment.what would it look like?

• Collect relevant data including segmentation variables and descriptor variables.

• Explore and preprocess the data.

• Split consumers into segments using segmentation variables and appropriate algorithms.

• Determine the key features of each segment profile.

• Describe each segment in detail.

• Evaluate the segments and select the target segment(s).

• Customize the marketing mix for the target segments.

• Evaluate the success of segmentation and monitor for changes.

• Market segmentation is a long term commitment that requires substantial changes and investments.

• Organizations must justify implementing segmentation through sufficient increase in sales to cover costs.

• Potential changes required include new product development, product modifications, pricing changes, distribution changes and communication changes.

• Internal organizational structure may also need to adjust to target different segments.

• Organizations need to organize around market segments to maximize segmentation benefits.

• The decision to pursue a market segmentation strategy must be made at the highest executive level and communicated throughout the organization.

• Barriers to successful market segmentation implementation include:

- Lack of leadership and commitment from senior management

- Insufficient resources allocated

- Organizational culture issues like resistance to change and lack of information sharing

- Short-term thinking

- Lack of training for staff on market segmentation fundamentals

• An organization's culture lacking a market or customer orientation can hinder market segmentation success.

• Step 2 involves specifying the ideal target segment for the organization.

• The organization must determine two sets of segment evaluation criteria: knock-out criteria and attractiveness criteria.

• Knock-out criteria are the essential, non-negotiable features of segments that the organization would consider targeting.

• Attractiveness criteria are used to evaluate the relative attractiveness of the remaining market segments that comply with the knock-out criteria.

• The literature proposes a wide array of possible segment evaluation criteria at different levels of detail.

• Knock-out criteria are essential criteria used to determine if market segments qualify to be assessed further.

• Some knock-out criteria include substantiality, measurability, accessibility and homogeneity of the segment.

• The segment must be homogeneous, meaning members of the segment must be similar to each other.

• The segment must be distinct, meaning members of the segment must be distinctly different from members of other segments.

• Attractiveness criteria represent a shopping list of criteria for the segmentation team to evaluate potential target segments.

• The segmentation team selects which attractiveness criteria to use and assesses the relative importance of each criterion.

• While knock-out criteria eliminate some segments, attractiveness criteria are used to determine the relative attractiveness of each segment.

• The target segment must be large enough to make customizing the marketing mix worthwhile.

• The target segment must match the organization's capabilities to satisfy members' needs.

• Members of the target segment must be identifiable in the marketplace.

• The target segment must be reachable so the organization can communicate the customized marketing mix.

• In addition to size and identifiability, attractiveness criteria like profit potential and growth rate help decide which segments to target.

• Segments are rated on how well they meet the attractiveness criteria, rather than strictly complying or not.

• Using a structured process when evaluating and selecting target markets, like a segment evaluation plot, is considered beneficial.

• A team of people should determine the segment attractiveness and organizational competitiveness values, using no more than six factors.

• It is recommended to have a core team of 2-3 people primarily in charge of the segmentation analysis who propose an initial solution and report to an advisory committee for discussion and modification.

• Representatives from different organizational units should be involved to provide different perspectives and ensure all units are stakeholders since the segmentation strategy will affect all units.

• The segment evaluation plot cannot be completed until after segments are identified. But selecting attractiveness criteria early ensures relevant information is captured during data collection and makes target segment selection easier later.

• The segmentation team should have a list of around six segment attractiveness criteria at the end, each with a weight indicating their relative importance. Team members typically distribute 100 points across the criteria which are then negotiated until agreement. Approval from the advisory committee is also recommended.

• Empirical data is used as the basis for both commonsense and data-driven market segmentation. It is used to identify and describe market segments.

• Segmentation variables refer to the variables used to split the sample into market segments.

• In commonsense segmentation, a single segmentation variable is typically used, such as gender.

• Descriptor variables describe the segments in detail and include demographics and media behavior.

• Data-driven segmentation uses multiple segmentation variables as the starting point to identify naturally existing or artificial market segments.

• The quality of empirical data is critical for developing a valid segmentation solution. It helps in assigning customers to correct segments, describing segments accurately, and developing customized products and promotions.

• Data for segmentation studies can come from survey studies, scanner data that records purchases, or experimental studies. Data that closely reflects actual consumer behavior is preferable.

• Before extracting segments, the organization must choose the segmentation criterion to use. The segmentation criterion can relate to a specific variable or construct, like benefits sought.

• The text discusses different types of market segmentation criteria. The most common ones are geographic, socio-demographic, psychographic and behavioral.

• The author argues that the simplest segmentation approach should be used. If demographic segmentation will work, use that. If geographic segmentation is sufficient, use that instead of more complex psychographic segmentation.

• Geographic segmentation divides customers based on their geographic location. While simple, it can be effective when targeting customers in specific regions.

• The advantages of geographic segmentation are that it is easy to assign customers to segments and target them using local communication channels.

• The disadvantage is that people in the same geographic area may not share other relevant characteristics for marketers.

• Geographic segmentation uses location as the segmentation variable. Though location makes segment membership easy to determine, it is rarely the actual reason for differences in product preferences.

• Socio-demographic segmentation uses criteria like age, gender, income and education. While it can be useful in some industries, socio-demographics often do not capture the real cause of product preferences.

• Studies show that socio-demographics explain only around 5% of the variance in consumer behavior. Experts argue that values, tastes and preferences are more influential in determining consumers' buying decisions.

• Psychographic segmentation groups people according to psychological criteria like beliefs, interests, preferences, and benefits sought when purchasing products.

• Benefit segmentation and lifestyle segmentation are popular forms of psychographic segmentation.

• Psychographic segmentation is complex due to the difficulty in finding a single characteristic that provides insight into the psychographic dimension. Multiple variables are often used.

• Psychographic segmentation is more reflective of the underlying reasons for differences in consumer behavior. For example, travel motives can identify tourists interested in cultural holidays.

• The disadvantages are the complexity and the reliance on accurate measures to capture the psychographic dimensions.

• Behavioral segmentation searches for similarities in actual or reported behaviors like purchase frequency, spending amounts, and information search behavior.

• The key advantage of behavioral segmentation is that actual behavior is used, which directly matches the segment of interest.

• Survey data is widely used for market segmentation as it is cheap and easy to collect. However, survey data can be affected by various biases which can negatively impact the segmentation solution.

• Carefully selecting the relevant variables to include as segmentation variables is critical for obtaining a high quality segmentation solution.

• Only variables relevant to the segmentation criterion should be included. Including unnecessary variables can make surveys longer and cause respondent fatigue, lowering data quality.

• Including unnecessary variables also increases the dimensionality of the segmentation problem, making it more difficult to extract market segments.

• Unnecessary "noisy" variables can prevent segmentation algorithms from identifying the correct solution.

STEP 7

• Step 7 involves describing segments using additional information available about segment members. It complements profiling segments based on the segmentation variables used to extract the segments.

• Good descriptions of market segments provide insight into the nature of segments and are essential for developing effective targeting and positioning strategies.

• Segment profiling investigates differences between segments with respect to the segmentation variables used to extract the segments.

• Segment description uses additional descriptor variables, such as demographic, psychographic, and behavior variables.

• Descriptor variables include age, gender, past travel behavior, preferred vacation activities, media use, information sources used during planning, and expenditure patterns.

• The example describes using travel motives as segmentation variables and then profiling and describing resulting segments using additional variables for a market segmentation analysis of the Australian travel market.

• Segment description reveals important information about the customers in a segment that can be used to develop a targeted marketing mix for that segment. Details like demographics, interests, media consumption habits, and contact preferences help shape the marketing strategy.

• Differences between market segments can be studied using descriptive statistics and visualizations or inferential statistics. Visualizations make segment descriptions more user-friendly.

• Visualizations like charts are useful for describing differences in nominal, ordinal and metric descriptor variables between market segments.

• Graphical visualizations simplify the interpretation of results for analysts and users. They also integrate information on statistical significance, avoiding overinterpretation of insignificant differences. Managers tend to prefer and find graphical displays more intuitive.

• The text discusses visualizing market segments using data from an Australian travel motives dataset.

• The dataset contains descriptor variables to describe market segments. Segment membership for respondents is stored in variable C6.

• The sizes of the 6 market segments range from 94 to 235 respondents.

• A cross-tabulation is created with segment membership as a categorical variable to analyze descriptor variables.

• A cross-tabulation of gender across segments suggests no huge gender differences.

• Stacked bar charts are used to visualize the cross-tabulation, showing segment sizes on the y-axis and gender within each bar.

• Mosaic plots are also used to visualize cross-tabulations, with the width indicating absolute segment size and column heights showing proportions within segments.

• The text compares a stacked bar chart and a mosaic plot for cross-tabulation of segment membership and gender.

• The number of segment members ranges from 0 to 250 for the 6 segments.

• The mosaic plot shows the proportion of men and women in each segment through the height of the rectangles.

• The width of the columns in the mosaic plot represents the total segment sizes.

• The area of each cell in the mosaic plot is proportional to the corresponding cell size in the table.

• Mosaic plots can visualize tables with more than two descriptor variables and incorporate elements of inferential statistics.

• Cell colors in mosaic plots can highlight where observed frequencies differ from expected frequencies based on the assumption that the variables are independent.

• A mosaic plot visualizes the relationship between two categorical variables using rectangles with areas proportional to the frequencies.

• Positive differences in the frequencies mean observed frequencies are higher than expected, and are colored blue. Negative differences mean observed frequencies are lower than expected and are colored red.

• By default in R, standardized Pearson residuals smaller than -4 are colored dark red, between -2 and -4 are light red, between -2 and 2 are white, between 2 and 4 are light blue, and above 4 are dark blue.

• The plot in Figure 9.2 shows that all cells are white, indicating that the six market segments extracted from the Australian travel motives data set do not significantly differ in gender distribution. The proportion of female and male tourists is approximately the same across segments.

• The figure shows a moderate association between tourist segment membership and income level. Tourists with higher incomes are more frequently part of segment 4, which values cultural offers and local people.

• Tourists with lower incomes are less likely to be part of segment 3, which seeks luxury, fun and entertainment.

• Tourists with very high incomes are underrepresented in segment 6, the nature loving segment.

• The figure points to a strong association between travel motives and moral obligation to protect the environment. Tourists with higher environmental moral obligation scores tend to have travel motives that value nature.

• The moral obligation score ranges from 1 to 5, with 1 being the lowest moral obligation and 5 being the highest moral obligation to protect the environment. This score is based on survey respondents' answers.

• The text discusses analyzing the moral obligation score in its original metric format versus recoding it into quarters.

• A mosaic plot illustrates the cross-tabulation of segment membership and stated moral obligation.

• Segment 3 members who seek entertainment have significantly lower stated moral obligation to act environmentally friendly compared to other segments.

• In contrast, segment 6 members who are motivated by nature have a positive association with high moral obligation to protect the environment.

• The text discusses using metric descriptor variables to visualize differences between market segments. The R packages lattice and ggplot2 can be used to create conditional plots divided into panels showing results for subsets of data.

• Histograms and box plots can be generated for different variables to compare the distributions across segments. This can show differences in variables like age and moral obligation scores between segments.

• While visually inspecting plots can reveal some differences between segments, statistical testing is needed to determine if the differences are statistically significant.

• Parallel box-and-whisker plots can incorporate statistical hypothesis testing, for example by making the width of the boxes proportional to the statistical significance.

• A parallel box-and-whisker plot was created to visualize the distribution of moral obligation to protect the environment across different market segments.

• The plot shows that segment 5 has the smallest size while segment 1 is the largest.

• Segment 6 has the highest moral obligation to protect the environment.

• The notches in the plot correspond to 95% confidence intervals for the medians. Non-overlapping notches between segments indicate a significant difference in their medians.

• There is a significant difference in moral obligation between members of segment 3 and segment 6 based on the non-overlapping notches between them.

• A parallel box-and-whisker plot was created to visualize the distribution of moral obligation to protect the environment across different market segments.

• The plot shows that segment 5 has the smallest size while segment 1 is the largest.

• Segment 6 has the highest moral obligation to protect the environment.

• The notches in the plot correspond to 95% confidence intervals for the medians. Non-overlapping notches between segments indicate a significant difference in their medians.

• There is a significant difference in moral obligation between members of segment 3 and segment 6 based on the non-overlapping notches between them.

• The text discusses using visualizations to describe market segments. It shows a parallel box-and-whisker plot of moral obligation to protect the environment by market segment.

• Most respondents state that they feel morally obliged to protect the environment, though some admit to not feeling a moral obligation.

• The text describes using a modified segment level stability across solutions (SLSA) plot to trace a metric descriptor variable over different market segmentation solutions. In this case, node color indicates each segment's mean moral obligation. Red indicates high obligation and gray indicates low obligation.

• The segment identified as a potentially attractive market segment consistently shows high moral obligation to protect the environment. The acquiescence segment also shows high obligation due to respondents' tendency to express agreement.

• Figure 9.9 shows the Segment Level Stability Across Solutions (SLSA) plot for the Australian travel motives data set with 3 to 8 segments. The nodes are colored by mean moral obligation values.

• Simple statistical tests can be used to formally test for differences in descriptor variables across market segments.

• The simplest way to test for differences is to run a series of independent tests for each variable of interest.

• Segment membership can be treated as a nominal summary statistic of the segmentation variables.

• Any test for association between a nominal variable and another variable can be used. This includes the chi-square test.

• The chi-square test can be used to formally test for significant differences in variables like gender distribution across market segments.

- The text discusses testing for segment differences using statistical tests like chi-squared test and ANOVA.

- The chi-squared test is used to test for differences in categorical variables like gender distribution between market segments. A small p-value indicates significant differences.

- The mosaic plot is used to visualize the chi-squared test results and identify which combinations occur more or less frequently.

- Boxplots are used to visualize differences in metric variables like age and spending between market segments.

- ANOVA is used to test for significant differences in the means of more than two groups, like the mean moral obligation values across market segments.

• The analysis of variance performs an F-test that compares the weighted variance between market segment means and the variance within market segments. Small F values support the null hypothesis that segment means are the same.

• The p-value below 0.05 indicates that we reject the null hypothesis that each segment has the same mean moral obligation. At least two market segments differ in their mean moral obligation.

• Presenting mean values of variables by segment in a table provides an overview of segment characteristics. The ANOVA p-values show if differences are significant.

• Pairwise comparisons between segments using t-tests identify which segments differ significantly.

• The results show that segment 1 does not differ significantly in mean moral obligation from segments 2, 3, and 4, but does differ significantly from segments 5 and 6.

• Segments 5 and 6 feel more morally obliged to protect the environment than segments 1, 2, 3 and 4.

• When performing a series of tests using the same data, p-values need to be adjusted for multiple testing to control the error rate.

• The text discusses different methods for correcting p-values for multiple testing including Bonferroni correction, Holm's method, and the false discovery rate procedure.

• As an alternative to pairwise t-tests, Tukey's honest significant differences can be plotted to compare group means.

- The plot shows the comparison of pairs of segments.

- The horizontal solid line depicts the confidence interval of the difference in mean values.

- If a confidence interval crosses the vertical line at 0, the difference is not significant.

- Segments 1,2,3 and 4 do not differ significantly from one another in moral obligation.

- Segments 5 and 6 are characterized by a significantly higher moral obligation to behave environmentally friendly than the other market segments, with the exception of segments 4 and 5 not differing significantly.

• Segment 4 sits between the low and high group, and does not display significant differences to segments 1–3 at the low end, and 5 at the high end.

• Another way of learning about market segments is to predict segment membership from descriptor variables using regression models and classification or supervised learning methods.

• Regression analysis assumes that a dependent variable y can be predicted using independent variables x1, . . . , xp.

• The basic regression model is the linear regression model which assumes that the relationship between y and the x variables is linear.

• Function lm() in R fits a linear regression model.

• The text describes using a regression model in R where the dependent variable AGE is indicated on the left side of the ~ and the independent variables are on the right side.

• Here only segment membership C6 as independent variable is used which is a categorical variable with six categories coded as a factor.

• For identifiability either the intercept β0 or one category needs to be dropped.

• When dropping the intercept, each estimated coefficient is equal to the mean age in that segment. Members of segment 5 have the lowest mean age of 39.4 years and segment 6 has the highest mean age of 49.6 years.

• When including the intercept $\beta 0$, the coefficient for segment 1 is dropped and its effect is captured by the intercept. The other coefficients indicate the mean age difference between segment 1 and the other segments.

• The text also explains that regression coefficients express how much the dependent variable changes if one independent variable changes while others remain constant.

• The linear regression model assumes changes caused by one independent variable are independent of the absolute level of all independent variables.

• Generalized linear models can accommodate a wider range of distributions for the dependent variable, which is important if the dependent variable is categorical.

•

• The linear regression model models the mean value of y given x1, ..., xp by the linear function: $E[y|x1, ..., xp] = \beta 0 + \beta 1x1 + ... + \beta pxp$.

• Generalized linear models are not limited to the normal distribution. They can use distributions like Bernoulli with y taking values 0 or 1.

• Generalized linear models use a link function g() to transform the mean value of y to an unlimited range indicated by η. This allows modeling with a linear function.

• η is called the linear predictor.

• Generalized linear models can use distributions like normal, Poisson, binomial, and multinomial for the dependent variable. Binomial and multinomial are needed for classification.

• Binary logistic regression assumes y follows a Bernoulli distribution and uses the logit link function.

• The logit link maps the success probability $\mu \in (0,1)$ to $(-\infty, \infty)$ using $\log(\mu/(1- \mu))$.

• The glm() function fits generalized linear models in R. The distribution and link function are specified using the family argument.

• The text describes fitting a binary logistic regression model to predict the likelihood of a consumer belonging to segment 3 based on their age and moral obligation score.

• The dependent variable is whether the consumer is in segment 3 (coded as a binary indicator) and the independent variables are age and obligation score.

• The model output provides the regression coefficients, model fit statistics like deviance and AIC, and information about the intercept.

• The intercept in logistic regression represents the linear predictor value when all independent variables equal 0.

• Transforming the intercept using the inverse logit function provides the predicted probability of belonging to segment 3 when all independent variables equal 0.

• The regression coefficients indicate how the linear predictor changes as the corresponding independent variable changes, which then impacts the odds of belonging to segment 3.

• The text discusses using odds ratios to analyze the probability of success versus failure. An odds ratio above 1 indicates success is more likely than failure.

• A binary logistic regression model was created (model.C63) to predict the probability of tourists belonging to segment 3 based on age and obligation scores.

• Increasing age was found to decrease the probability of belonging to segment 3, though age did not have a statistically significant effect based on the confidence bands.

• Increasing obligation scores were also found to decrease the probability of belonging to segment 3. Tourists with the highest obligation scores had around half the probability of those with the lowest scores, and this effect was statistically significant.

• The package "effects" was used to calculate and plot the predicted probabilities for different levels of the independent variables while holding other variables constant at their averages.

• The output contains the estimated coefficients, their standard errors, test statistics of a z-test, and the associated p-values. The z-test compares the fitted model to a model where the regression coefficient is set to 0.

•AGE is not significant based on the z-test. Dropping AGE from the model would not significantly decrease model fit.

• The variable OBLIGATION2 has a significant impact based on the analysis of deviance test. Dropping OBLIGATION2 would significantly reduce model fit. So moral obligation is useful in predicting segment 3.

• Additional independent variables can be added to the binary logistic regression model.

• The output contains the estimated coefficients, their standard errors, test statistics of a z-test, and the associated p-values for a binary logistic regression model.

• The AGE variable is not statistically significant based on the z-test. Excluding AGE from the model likely would not meaningfully impact model fit.

• The OBLIGATION2 variable has a significant impact according to the analysis of deviance test. Removing OBLIGATION2 would likely significantly degrade model fit. So moral obligation helps predict segment 3.

• Additional predictor variables can be included in the binary logistic regression model.

• The text discusses using market segmentation to predict consumer behavior and create targeted marketing strategies.

• Stepwise regression and multinomial logistic regression models are fitted to predict market segments.

• Multinomial logistic regression can predict multiple segments simultaneously, while stepwise regression predicts segments separately.

• The multinomial logistic regression model shows that age and obligation to vacation have some effect on predicting market segments.

• The regression coefficients indicate the change in log odds of a consumer belonging to that segment as the independent variables change.

• Classification and regression trees are a supervised learning technique that can perform variable selection, have an easy interpretation and incorporate interaction effects.

• Trees work by splitting consumers into groups based on independent variables to achieve high purity in the dependent variable.

• The resulting tree shows nodes that emerge from each split, with terminal nodes that are not split further. Segment membership is predicted based on the terminal node.

• Tree algorithms differ in the type of splits, selection criteria, stopping criteria and predictions.

• The rpart and partykit R packages implement tree algorithms.

STEP 10

• Market segmentation analysis is an ongoing process that requires continuous evaluation and monitoring.

• After implementing the segmentation strategy, two tasks should be performed:

- Evaluate the effectiveness of the segmentation strategy to determine if it achieved the desired benefits.

- Monitor the market continuously since it is constantly changing.

• The aim of evaluating the segmentation strategy is to determine if customizing marketing mix for target segments achieved expected benefits for the organization.

• In the short term, benefits include increased sales and profits. In the long term, benefits include increased market share and brand equity.

• An ongoing monitoring process should be established to detect changes in the target segment size and characteristics as well as actions of competitors.

• The primary desired outcome for most organizations is increased profit or for non-profits it could be donations or volunteers.

• Segment performance can be monitored continuously to assess the segmentation strategy.

• Segment membership is found to be unstable over time with less than one-third of customers remaining in the same segment after two years.

• Changes in segment size and characteristics may require modifications to the segmentation strategy and marketing mix.

• Segment hopping can occur due to different consumption situations, seeking variety, and promotional offers.